

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

1 nucleic - nucleic search, using sw model  
in on: February 1, 2004, 19:35:31 ; Search time 12453 Seconds  
(without alignments)  
11484.787 Million cell updates/sec  
tble: US-09-042-460-1  
arfect score: 3496  
quence: 1 GAATTCGGGTGGAGGCC.....CCGAGTCGTACCAAGCTT 3496  
oring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
sarched: 2888711 seqs, 2045481386 residues  
otal number of hits satisfying chosen parameters: 5777422  
inimum DB seq length: 0  
aximum DB seq length: 2000000000  
ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.on.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg\_hum.\*
- 31: em.htg\_inv.\*
- 32: em.htg\_other.\*
- 33: em.htg\_mus.\*
- 34: em.htg\_pln.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_mam.\*
- 37: em.htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3426	98.0	3426	10	AF051911	AF051911 Mus muscu
2	3367.4	96.3	3369	10	AF073311	AF073311 Mus muscu
3	2578.6	73.8	4170	10	AF149012	AF149012 Mesocric
4	1758	50.2	1758	10	AF247818	AF247818 Rattus no
5	1603.6	45.3	8660	12	AF043739	AF043739 Synthetic
6	1600.8	45.8	4015	6	AR104587	AR104587 Sequence
7	1600.8	45.8	4015	6	ARI75848	ARI75848 Sequence
8	1600.8	45.8	4015	6	AR182221	AR182221 Sequence
9	1600.8	45.8	4015	6	AR224455	AR224455 Sequence
10	1600.8	45.8	4015	6	AR226390	AR226390 Sequence
11	1600.8	45.8	4015	6	AR243328	AR243328 Sequence
12	1600.8	45.8	4015	6	AR263555	AR263555 Sequence
13	1600.8	45.8	4015	6	AR265996	AR265996 Sequence
14	1600.8	45.8	4015	6	AX019310	AX019310 Sequence
15	1600.8	45.8	4015	6	AX133979	AX133979 Sequence
16	1600.8	45.8	4015	6	AX552695	AX552695 Sequence
17	1600.8	45.8	4015	6	BD011044	BD011044 Human tel
18	1600.8	45.8	4015	6	BD015832	BD015832 Quantitat
19	1600.8	45.8	4015	6	BD082985	BD082985 Method fo
20	1600.8	45.8	4015	6	BD131727	BD131727 Method fo
21	1600.8	45.8	4015	6	E36793	E36793 Human telom
22	1600.8	45.8	4015	9	AF015950	AF015950 Homo sapi
23	1600.8	45.8	4027	9	AF018167	AF018167 Homo sapi
24	1600.8	45.8	4042	6	AX001446	AX001446 Sequence
25	1600.8	45.8	4042	6	AX003121	AX003121 Sequence
26	1600.8	45.8	4042	6	BD136185	BD136185 Human tel
27	1600.8	45.8	4070	6	AX331846	AX331846 Sequence
28	1596	45.7	3399	6	AX481414	AX481414 Sequence
29	1596	45.7	4037	6	BD011070	BD011070 Human tel
30	1594.6	45.6	3396	6	BD091553	BD091553 Adult bon
31	1594.6	45.6	3396	6	BD094749	BD094749 The cell
32	1594.6	45.6	3396	6	BD096291	BD096291 Cells cap
33	1593.6	45.6	13766	6	AX553919	AX553919 Sequence
34	1593.6	45.6	8742	6	AX553924	AX553924 Sequence
35	1584.8	45.3	8742	6	AR104586	AR104586 Sequence
36	1520.2	43.5	4029	6	AR175847	AR175847 Human tel
37	1520.2	43.5	4029	6	BD011069	BD011069 Human tel
38	1520.2	43.5	4029	6	E36818	E36818 Human telom
39	1520.2	43.5	4029	6	BD123426	BD123426 Transgeni
40	1462.6	41.8	7498	10	AF121949	AF121949 Mus muscu
41	1462.6	41.8	216363	2	AC131076	AC131076 Mus muscu
42	1462.6	41.8	216363	2	AC131076	AC131076 Mus muscu
43	1323.2	37.8	3855	6	AR243330	AR243330 Sequence
44	1323.2	37.8	3855	6	BD011046	BD011046 Human tel
45	1323.2	37.8	3855	6	E36795	E36795 Human telom

ALIGNMENTS

RESULT 1  
AF051911  
LOCUS 3426 bp mRNA linear ROD 01-DEC-2000  
DEFINITION Mus musculus telomerase reverse transcriptase mRNA, complete cds.  
ACCESSION AF051911  
VERSION AF051911.1 GI:3005591  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 3426)  
REFERENCE  
AUTHORS Greenberg,R.A., Allsopp,R.C., Chin,L., Morin,G.B. and Depinho,R.A.  
TITLE Expression of mouse telomerase reverse transcriptase during  
development, differentiation and proliferation

JOURNAL Oncogene 16 (13), 1723-1730 (1998)  
 3DLINE 98241176  
 PUBMED 9582020  
 ERENCE 2 (bases 1 to 3426)  
 AUTHORS Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-MAR-1998) Microbiology and Immunology, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA

TURES  
 source  
 1. 3426  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="13"  
 /map="Mdb offset position 40; 40 cm from centromere"  
 30. 3398  
 /note="mPRT"  
 /codon\_start=1  
 /product="telomerase reverse transcriptase"  
 /protein\_id="AAC09323.1"  
 /db\_xref="GI:300592"  
 /translation="MTRAPCPAVSLRLSRYSREVPPLATFVRLGPEGRRLVQPGDP  
 KYETVAOCLVCMWQSPADLSFHOSSLSKELVARVORLCSERNVLAEGE  
 LNAEAGGPMAFTSVRSYLPNVVETLAVSGAMLLLSRVGDDLLVLLHCLAYL  
 LVFPCAYQCCSPFLICATIDIPSVASISRETRVGRNFTNLRLPQIIKSSRSP  
 APFLPALPGRTRHLSLTSVSPAKARCPVFRVEGPHRQVLPTPGSKSWPSP  
 ARSPVPTAKDLSSKGVSDLSUSGVCCCKHPSSTLSLSPERQALFQIRFETRH  
 FLSRGDQRLNPSFLNLSQNLNLTGARLVEILFSGSPRTSGPLCTRHLSSRYW  
 QMLPFLQLLVNHASQVYLLRHSCFRPTANQQVTDALNTPHMLDLRLHSSPWQ  
 VYGLFRACLVKVSASLGRNRRFPNKKFISLKYKLSLQELMWMKMKVEDCH  
 WLRSSPKDRVPAAEHRLRLATFLFMDLVVOLLRSFFVIFESTFOKNRLLFY  
 KSVWKLQIGVROHLERLRLSLOEVARHODTWLAMPICRLRFIPKNGRLPIV  
 NMSVSNLGRALGRKQAFQRLKTLFSLNRYNTRKHPLMGSSVLGMDIYRTWA  
 FVLRLVADOTRMVFEVADVTGADALPQGLKLVWVWAMIRSESTYCIQYAVR  
 DSQGVHFRQVITLSDLPYMGQFLKHLQSDASALNSVIVQSISWNESSSL  
 PFFFLHLSHVSVCITCQGIPOGSLTSLICSLGDMENKLVSVVQRDGLL  
 LRPFDFLLVTHLOQATFLSTLVHGVPEYGMINLQTVNVPVPEPTLGGAPYQ  
 LRPVDFPFCWGLDQTLFVFCYSYATKISLTSQSVFPAKTRNKLKSLVR  
 LKCHGLFLDQVNSLQVINCINIKYIFLLQAYRHACVIOLEPDRVRKHLFFGLIIS  
 SQSCCYAILKVNPMKTKASGFPPEAAHWCYQAFLLKLAHSHVYKCLGLPLRT  
 AQLKRLKPLPMTWILKAAADPALSTDFQILD"  
 746 a 979 c 907 g 794 t

ASE COUNT  
 IGIN

Query Match 98.0%; Score 3426; DB 10; Length 3426;  
 Best Local Similarity 100.0%; Fred. No. 0;  
 Matches 3426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 10 GTGGAGGCCATCCCGGCTTTGAGCAACATGACCGCGCTCTCTGTTGCGCGGTCG 69  
 b 1 GTGGAGGCCATCCCGGCTTTGAGCAACATGACCGCGCTCTCTGTTGCGCGGTCG 60  
 Y 70 GCTCTCTGCTGCGAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGCGCGGCG 129  
 b 61 GCTCTCTGCTGCGAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGCGCGGCG 120  
 Y 130 TGGGGCCGAGGCGCGGCTTTGCAACCGGGGACCCGAGATCTACCGCACTTTGG 189  
 b 121 TGGGGCCGAGGCGCGGCTTTGCAACCGGGGACCCGAGATCTACCGCACTTTGG 180  
 Y 190 TTGCCCAATCCCTAGTGTGATGCACTGGGGCTCACAGCCTCCACCTGCGGACCTTTCT 249  
 b 181 TTGCCCAATCCCTAGTGTGATGCACTGGGGCTCACAGCCTCCACCTGCGGACCTTTCT 240  
 Y 250 TCCACCAAGTGTCTATCCCTGAAGAGCTGTGGCGAGGTTGTGCAAGACTCTGCGAGC 309  
 b 241 TCCACCAAGTGTCTATCCCTGAAGAGCTGTGGCGAGGTTGTGCAAGACTCTGCGAGC 300  
 Y 310 GCAACGAGAGAAACGTCGTGCTTTGGCTTTGAGCTGCTTTAACAGGCCGAGGCGGCG 369  
 b 301 GCAACGAGAGAAACGTCGTGCTTTGGCTTTGAGCTGCTTTAACAGGCCGAGGCGGCG 360

QY 370 CTCCTCATGGCTTCACTAGTAGCGTAGCTACTTGGCCAAACACTGTTATTGAGACCC 429  
 Db 361 CTCCTCATGGCTTCACTAGTAGCGTAGCTACTTGGCCAAACACTGTTATTGAGACCC 420  
 QY 430 TCGGTGTCAGTGGGTCAGTGGTACTTCTAGCCGAGTGGGCGAGCACTGCTGTCT 489  
 Db 421 TCGGTGTCAGTGGGTCAGTGGTACTTCTAGCCGAGTGGGCGAGCACTGCTGTCT 480  
 QY 490 ACTGCTGCGCACACTGCTGCTTTATCTTCTGTCGCCGCCAGTGTGCTTACCAAGTGT 549  
 Db 481 ACTGCTGCGCACACTGCTGCTTTATCTTCTGTCGCCGCCAGTGTGCTTACCAAGTGT 540  
 QY 550 GTGGGTCTCCCTCTGTACCAATTTGTGCCACACAGGATATCTGGCCCTCTGTGTCGCTA 609  
 Db 541 GTGGGTCTCCCTCTGTACCAATTTGTGCCACACAGGATATCTGGCCCTCTGTGTCGCTA 600  
 QY 610 GTTACAGGCCACCCGACCCGTCGAGGAGGAAATTTCACTAACCTTAGGTCTTACAACAGA 669  
 Db 601 GTTACAGGCCACCCGACCCGTCGAGGAGGAAATTTCACTAACCTTAGGTCTTACAACAGA 660  
 QY 670 TCAAGAGCAGTAGTGGCGAGGAGACCCGAAACCCCTGGCTTGCCTATCTCGAGGTACAA 729  
 Db 661 TCAAGAGCAGTAGTGGCGAGGAGACCCGAAACCCCTGGCTTGCCTATCTCGAGGTACAA 720  
 QY 730 AGAGGCATCTGAGTCTCACAGTACAGTGTGCTTACGTAAAGAGCCAGATGCTATC 789  
 Db 721 AGAGGCATCTGAGTCTCACAGTACAGTGTGCTTACGTAAAGAGCCAGATGCTATC 780  
 QY 790 CTGTCCCGAGAGTGGAGGAGGAGCCCAACAGGAGGTGCTACCAACCCCTCAGGCAAAAT 849  
 Db 781 CTGTCCCGAGAGTGGAGGAGGAGCCCAACAGGAGGTGCTACCAACCCCTCAGGCAAAAT 840  
 QY 850 CATGGGTGCCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 909  
 Db 841 CATGGGTGCCAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 910 TTAAGAGAAAGGTGTCTGACCTGAGTCTCTCTGCGGTGCGGTGTGCTGCTGCTGCTGCT 969  
 Db 901 TTAAGAGAAAGGTGTCTGACCTGAGTCTCTCTGCGGTGCGGTGTGCTGCTGCTGCTGCT 960  
 QY 970 GCTCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029  
 Db 961 GCTCCACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
 QY 1030 AGACGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1089  
 Db 1021 AGACGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
 QY 1090 TACTCAGCAACCTCCAGCCTTAACTTGAATGGGCGCAGAGACTGGTGGAGATCATCTTTC 1149  
 Db 1081 TACTCAGCAACCTCCAGCCTTAACTTGAATGGGCGCAGAGACTGGTGGAGATCATCTTTC 1140  
 QY 1150 TGGGTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCGTCTATCGGTGCTAT 1209  
 Db 1141 TGGGTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCGTCTATCGGTGCTAT 1200  
 QY 1210 ACTGGCAGATGGCGGCCCTGTTTCAACAGCTGCTGTTGAACATGCGAGAGTGGCAATATG 1269  
 Db 1201 ACTGGCAGATGGCGGCCCTGTTTCAACAGCTGCTGTTGAACATGCGAGAGTGGCAATATG 1260  
 QY 1270 TCAGACTCTCAGGTCACTTTCAGGTGTTTGAACAGCAACCAACAGGTGACAGATGCTCT 1329  
 Db 1261 TCAGACTCTCAGGTCACTTTCAGGTGTTTGAACAGCAACCAACAGGTGACAGATGCTCT 1320  
 QY 1330 TGAACACCGCCACCGACCTCTCATGGATTTCTCCGCTGCACAGCAGTCCCTGGCAGG 1389  
 Db 1321 TGAACACCGCCACCGACCTCTCATGGATTTCTCCGCTGCACAGCAGTCCCTGGCAGG 1380  
 QY 1390 TATATGTTTCTTGGGCTGCTGCTCTGCAAGGTGGTGTCTGCTAGTCTCTGGGTACCA 1449  
 Db 1381 TATATGTTTCTTGGGCTGCTGCTCTGCAAGGTGGTGTCTGCTAGTCTCTGGGTACCA 1440  
 QY 1450 GGCACATGAGCGCGCTTCTTTAAGAACTTAAAGAGTTCATCTCTGTTGGGAAATACG 1509



1441	GGCAAAATGAGCGCGCTTCTTTAAGAACTTAAAGAAAGTTCACTCTCGTTGGGGAATAAG	1500	2521	CCACCTGCTCTGCACTGTTGTTTCGGAGACATGGAGACAAAGCTGTTTGTCTGAGGTGC	2580
1510	GCAGCTATCACTGCAGGAACATGATGCGAGACATGAAGTAGAGATTGCCACTGGCTCC	1569	2590	AGCGGATGGGTTGCTTTTACGTTTGTGTGATGATTTCTGTGTGAGACGCTCACTTGG	2649
1501	GCAGCTATCACTGCAGGAACATGATGCGAGACATGAAGTAGAGATTGCCACTGGCTCC	1560	2581	AGCGGATGGGTTGCTTTTACGTTTGTGTGATGATTTCTGTGTGAGACGCTCACTTGG	2640
1570	GCAGAGCCCGGGGAAGCACTGTGCTCCCGCTGCAGAGCACCGTCTGAGGGAGAGATCC	1629	2650	ACCAAGCMAAAACCTTCTCAGCACCTGCTCATGGCGTCTCTAGTAGTATGGGTGCATGA	2709
1561	GCAGAGCCCGGGGAAGCACTGTGCTCCCGCTGCAGAGCACCGTCTGAGGGAGAGATCC	1620	2641	ACCAAGCMAAAACCTTCTCAGCACCTGCTCATGGCGTCTCTAGTAGTATGGGTGCATGA	2700
1630	TGGCTACGTTCTGTTCTGGCTGATGGACACATAGTGTATCAGCTGCTTATAGTCACTTCT	1689	2710	TAAACTTGCAGAAAGACAGTGGTGAACTTCCCTGTGTGAGCCTGGTACCCCTGGGTFGGCAG	2769
1621	TGGCTACGTTCTGTTCTGGCTGATGGACACATAGTGTATCAGCTGCTTATAGTCACTTCT	1680	2701	TAAACTTGCAGAAAGACAGTGGTGAACTTCCCTGTGTGAGCCTGGTACCCCTGGGTFGGCAG	2760
1690	TTTATCATCACAGAGACATTTCCAGAAAGACAGGCTCTTTCTTACCCGTAAAGATGTGT	1749	2770	CTCCATACAGCTGCCTGCTCACTGCTGTTTCCCTGGTGTGGCTTGTCTGTGACACTC	2829
1681	TTTATCATCACAGAGACATTTCCAGAAAGACAGGCTCTTTCTTACCCGTAAAGATGTGT	1740	2761	CTCCATACAGCTGCCTGCTCACTGCTGTTTCCCTGGTGTGGCTTGTCTGTGACACTC	2820
1750	GGAGCAAGCTGCGAGCATGTGAGTTCAGCAACACTTGCAGAGATGGCGTACGGGAGC	1809	2830	AGACTTGGAGGTTCTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCC	2889
1741	GGAGCAAGCTGCGAGCATGTGAGTTCAGCAACACTTGCAGAGATGGCGTACGGGAGC	1800	2821	AGACTTGGAGGTTCTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCC	2880
1810	TGTCACAGAGGAGTTCAGGATCACAGCAACCTTGGCTAGCCATGCCCATCTGCAGAC	1869	2890	TCACCTTCCAGAGTGTCTTCAAAGCTGGGAAGACCATGCGGAACAAGCTCCTGTCCGTCT	2949
1801	TGTCACAGAGGAGTTCAGGATCACAGCAACCTTGGCTAGCCATGCCCATCTGCAGAC	1860	2881	TCACCTTCCAGAGTGTCTTCAAAGCTGGGAAGACCATGCGGAACAAGCTCCTGTCCGTCT	2940
1870	TGGCTTCACTCCCAAGCCCAACGGCTGCGGCCCATTTGTGAACATGATTTATAGCATGG	1929	2950	TGGGTTGAAGTGTCAAGGTTCTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCT	3009
1861	TGGCTTCACTCCCAAGCCCAACGGCTGCGGCCCATTTGTGAACATGATTTATAGCATGG	1920	2941	TGCGTTGAAGTGTCAAGGTTCTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCT	3000
1930	GTACCAAGGCTTTGGCAGAGGAAGCAAGCCAGCATTTTCAACCAGCGTCTCAAGACTC	1989	3010	GCATCAATATATACAGATCTTCTGCTTCAAGGCTTCAAGGTTCCATGCAATGTGTGATTC	3069
1921	GTACCAAGGCTTTGGCAGAGGAAGCAAGCCAGCATTTTCAACCAGCGTCTCAAGACTC	1980	3001	GCATCAATATATACAGATCTTCTGCTTCAAGGCTTCAAGGTTCCATGCAATGTGTGATTC	3060
1990	TCTTCAGCATGCTCAACTATGAGCGCAAAACATCCTCACTTATGGGGTCTTCTGTAC	2049	3070	AGCTTCCCTTTGACCGCGTGTAGGAAGAACCTTCACTTCTTCTGGGCTCATCTCTCCA	3129
1981	TCTTCAGCATGCTCAACTATGAGCGCAAAACATCCTCACTTATGGGGTCTTCTGTAC	2040	3061	AGCTTCCCTTTGACCGCGTGTAGGAAGAACCTTCACTTCTTCTGGGCTCATCTCTCCA	3120
2050	TGGGTATGAATGATCTACAGACCTGGGGGCTTTGTGCTGGTGTGGCTCTGG	2109	3130	GCCAAGCATCTGCTGCTATGCTATCCTGAAGGTCAAGATCCAGGAATGACACTAAAGG	3189
2041	TGGGTATGAATGATCTACAGACCTGGGGGCTTTGTGCTGGTGTGGCTCTGG	2100	3121	GCCAAGCATCTGCTGCTATGCTATCCTGAAGGTCAAGATCCAGGAATGACACTAAAGG	3180
2110	ACCAGACACCCAGGATGATCTTTGTTAAGGCAGATGTGACCGGGGCTTATGATCCATCC	2169	3190	CCTCTGGCTCTTCTCTCTGAGCGCACATTTGGCTCTGCTACCGGCTTCTCTGCTCA	3249
2101	ACCAGACACCCAGGATGATCTTTGTTAAGGCAGATGTGACCGGGGCTTATGATCCATCC	2160	3181	CCTCTGGCTCTTCTCTCTGAGCGCACATTTGGCTCTGCTACCGGCTTCTCTGCTCA	3240
2170	CCAGGGTAAAGTGTGGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACT	2229	3250	AGCTGGCTGCTCATTTCTGTCTATCTACAAATGTCTCTCTGGGACCTCTGAGGACAGCCAAA	3309
2161	CCAGGGTAAAGTGTGGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACT	2220	3241	AGCTGGCTGCTCATTTCTGTCTATCTACAAATGTCTCTCTGGGACCTCTGAGGACAGCCAAA	3300
2230	GTATCCGCCAGTATGCACTGGTCCGAGAGATAGCAAGCCCAAGTCCCAAGTCTCTTTA	2289	3310	AACCTGTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGCTGCAAGCTGACC	3369
2221	GTATCCGCCAGTATGCACTGGTCCGAGAGATAGCAAGCCCAAGTCCCAAGTCTCTTTA	2280	3301	AACCTGTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGCTGCAAGCTGACC	3360
2290	GGAGACAGGTCAACACCTCTCTGACCTCCAGGCCATACATGGGCCAGTTCCTTAAAGCATC	2349	3370	CAGCCCTAAGCACAGACTTTTCAGACCAATTTTGGACTAAACCTGTCTCTTCCGTAGATG	3429
2281	GGAGACAGGTCAACACCTCTCTGACCTCCAGGCCATACATGGGCCAGTTCCTTAAAGCATC	2340	3361	CAGCCCTAAGCACAGACTTTTCAGACCAATTTTGGACTAAACCTGTCTCTTCCGTAGATG	3420
2350	TGCAGGATTCAGATGCGAGTGCATGAGGAACCTCCGTTGTCATCGAGCAGAGCATCTTA	2409	3430	AACATG 3435	
2341	TGCAGGATTCAGATGCGAGTGCATGAGGAACCTCCGTTGTCATCGAGCAGAGCATCTTA	2400	3421	AACATG 3426	
2410	TGAAATGAGCAGCAGCGCTGTTTGAATCTTCTTCTGCACTTCTCTGGTCAAGTGTGG	2469			
2401	TGAATGAGCAGCAGCAGCGCTGTTTGAATCTTCTTCTGCACTTCTCTGGTCAAGTGTGG	2460			
2470	TAAAGATTGTGACAGGTGCTATACGAGTGCAGGCGATCCCGCAGGGCTCCAGCCTAT	2529			
2461	TAAAGATTGTGACAGGTGCTATACGAGTGCAGGCGATCCCGCAGGGCTCCAGCCTAT	2520			
2530	CCACCTGCTCTGCACTGCTGTTTTCGGAGACATGGAGAACAAAGCTGTTTGTCTGAGGTGC	2589			

RESULT 2  
AF073311  
LQUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AF073311 3369 bp mRNA linear ROD 09-SEP-1998  
Mae musculus telomerase catalytic subunit mRNA, complete cds.

AF073311  
AF073311.1 GI:3551846

• *Mus musculus* ('house mouse')

[illegible]



1479	TTAAGAGGTTCTATCTCGTTGGGAAATACGCAAGCTATCACTCGAGAACTGATGG	1538	2521	GACATGGAGAACACAGCTGTTTCTGCTGAGGTGACGGGATGGGTTGCTTTACGTTTGT	2580
1441	TTAAGAAAGTTCTATCTCGTTGGGAAATACGCAAGCTATCACTCGAGAACTGATGG	1500	Qy	GATGACTTTTCTGTTGGTGAACCTTACCTTGGACCAAGCAAAACCTTCTCTCAGACACCTG	2678
1539	AAGATGAAGTACAGGATTCGCACTGGCTCCGACAGACCGGGGAAGACCGTGTCC	1598	Db	GATGACTTTTCTGTTGGTGAACCTTACCTTGGACCAAGCAAAACCTTCTCTCAGACACCTG	2640
1501	AAGATGAAGTACAGGATTCGCACTGGCTCCGACAGACCGGGGAAGACCGTGTCC	1560	Qy	GTCCATGGCGTTCTGAGTATGGGTGCAATGAATTTGACAGAGACAGTGGTGAACCTTC	2738
1599	GCTGCAGAGCACCGTCTGAGGGAGAGGATCTCTGGTACGTTCTCTGTTCTGGCTGATGAC	1658	Db	GTCCATGGCGTTCTGAGTATGGGTGCAATGAATTTGACAGAGACAGTGGTGAACCTTC	2700
1561	GCTGCAGAGCACCGTCTGAGGGAGAGGATCTCTGGTACGTTCTCTGTTCTGGCTGATGAC	1620	Qy	CTGTGTGGAGCCCTGGTACCCCTGGGTGGTGCAGCTCCATACACAGCTGCCTCACTGCCTG	2798
1659	ACATACGTGTACAGCTGTCTTGGTCACTTCTTATCATCACAGAGACACATCCAGAG	1718	Db	CTGTGTGGAGCCCTGGTACCCCTGGGTGGTGCAGCTCCATACACAGCTGCCTCACTGCCTG	2760
1621	ACATACGTGTACAGCTGTCTTGGTCACTTCTTATCATCACAGAGACACATCCAGAG	1680	Qy	TTTCCCTGTGTGGCTTGTCTGTGACACTCAGACTTTTGGAGGTGTTCTGTGACTACTCA	2858
1719	AACAGGCTCTTCTTCTACCGTAAAGAGTGTGTGAGCAAGCTGCAGAGCAATTGGAGTCA	1778	Db	TTTCCCTGTGTGGCTTGTCTGTGACACTCAGACTTTTGGAGGTGTTCTGTGACTACTCA	2820
1681	AACAGGCTCTTCTTCTACCGTAAAGAGTGTGTGAGCAAGCTGCAGAGCAATTGGAGTCA	1740	Qy	GGTTATGCCACAGACCTCAATTAAAGACGAGCCTCACCTTCCAGAGTGTCTTCAAAGCTGGG	2918
1779	CAACACCTTGAGAGATGCGGCTACCGGAGCTGTCAAGAGAGAGTCAAGGATCACACG	1838	Db	GGTTATGCCACAGACCTCAATTAAAGACGAGCCTCACCTTCCAGAGTGTCTTCAAAGCTGGG	2880
1741	CAACACCTTGAGAGATGCGGCTACCGGAGCTGTCAAGAGAGAGTCAAGGATCACACG	1800	Qy	AAGACCATGCGGAAACAGCTCCTGTGCGTCTTGGGTGAAAGTGTCAAGGCTATTTCTTA	2940
1839	GACACCTGGCTAGCATTGCCATCTGCAGACTGGCTTATCCCAAGCCCAAGCGCCTG	1898	Db	AAGACCATGCGGAAACAGCTCCTGTGCGTCTTGGGTGAAAGTGTCAAGGCTATTTCTTA	2940
1801	GACACCTGGCTAGCATTGCCATCTGCAGACTGGCTTATCCCAAGCCCAAGCGCCTG	1860	Qy	GACTTCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTT	3038
1899	CGGCCCATTTGTGAATGAGTTATAGCATGGGTACAGAGCTTTGGGCAAGAGAGGAG	1958	Db	GACTTCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTT	3000
1861	CGGCCCATTTGTGAATGAGTTATAGCATGGGTACAGAGCTTTGGGCAAGAGAGGAG	1920	Qy	CAGGCTTACAGGTTCCATGCATGTGTGATTCAAGTTCCTTTTGACAGCGTGTAGGAG	3098
1959	GCCAGCATTTCCCGCAGCTCTCAAGCTCTCTTCAGCATGCTCAACTATGAGCGGACA	2018	Db	CAGGCTTACAGGTTCCATGCATGTGTGATTCAAGTTCCTTTTGACAGCGTGTAGGAG	3060
1921	GCCAGCATTTCCCGCAGCTCTCAAGCTCTCTTCAGCATGCTCAACTATGAGCGGACA	1980	Qy	AACCTCACATTTCTTCTGGGATCATCTCCAGCAAGCATCCTGCTGTATCTATCTCTG	3158
2019	AAACATCTCTACCTTATGGGGTCTTCTGTACTGGGTATGAATGACATCTACAGGACCTG	2078	Db	AACCTCACATTTCTTCTGGGATCATCTCCAGCAAGCATCCTGCTGTATCTATCTCTG	3120
1981	AAACATCTCTACCTTATGGGGTCTTCTGTACTGGGTATGAATGACATCTACAGGACCTG	2040	Qy	AAGCTCAAGAAATCCAGGAATGACACTAAAGGGCTCTGGCTCTCTCTCTCTGAAAGCGCA	3218
2079	CGGGCTTTGTGCTGCTGCTGCTCTGAGACAGACACCCAGGATGACTTTGTTAG	2138	Db	AAGCTCAAGAAATCCAGGAATGACACTAAAGGGCTCTGGCTCTCTCTCTGAAAGCGCA	3180
2041	CGGGCTTTGTGCTGCTGCTGCTGAGACAGACACCCAGGATGACTTTGTTAG	2100	Qy	CATTGGCTCTGCTACAGGCTTCTGCTCAAGCTTCCCTGCTGCTCATCTGTCACTACAAA	3278
2139	GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGTAAGCTGGTGAAGTGTGTGCC	2198	Db	CATTGGCTCTGCTACAGGCTTCTGCTCAAGCTTCCCTGCTGCTCATCTGTCACTACAAA	3240
2101	GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGTAAGCTGGTGAAGTGTGTGCC	2160	Qy	TGCTCTCTGGGACCTCTGAGGACAGCCCAAAACCTGCTGTGCCGGAAGCTCCAGAGCG	3338
2199	AATATGATCAGGCACTCGGAGAGACGTAATGATCCCGCAGATGAGTGGTCCGAGA	2258	Db	TGCTCTCTGGGACCTCTGAGGACAGCCCAAAACCTGCTGTGCCGGAAGCTCCAGAGCG	3300
2161	AATATGATCAGGCACTCGGAGAGACGTAATGATCCCGCAGATGAGTGGTCCGAGA	2220	Qy	ACAATGACCATCTTAAAGCTGAGCTGACCCAGCCCTTAAGCAGACAGCTTTCCAGACCAT	3398
2259	GATAGCAAGGCAAGTCCCAAGTCTTTAGGAGACAGTCAACACCTCTCTGACCTC	2318	Db	ACAATGACCATCTTAAAGCTGAGCTGACCCAGCCCTTAAGCAGACAGCTTTCCAGACCAT	3360
2221	GATAGCAAGGCAAGTCCCAAGTCTTTAGGAGACAGTCAACACCTCTCTGACCTC	2280	Qy	TTGGACTAA 3407	
2319	CAGCCATACATGGCCAGTTCCTTAAGCATCTGAGGATTCAGATGCACTGAGG	2378	Db	TTGGACTAA 3369	
2281	CAGCCATACATGGCCAGTTCCTTAAGCATCTGAGGATTCAGATGCACTGAGG	2340	RESULT 3		
2379	AACTCGGTGTATCGAGAGAGATCTCTATGAATGAGAGCAGCAGCGCTTTTGAC	2438	AF149012		
2341	AACTCGGTGTATCGAGAGAGATCTCTATGAATGAGAGCAGCAGCGCTTTTGAC	2400	LOCUS	4170 bp mRNA linear ROD 21-FEB-2001	
2439	TTCTTCTCTGCACTTCTGCTGCTCAGAGTGTCTGAAGATTTGTTGATGAGTGTG	2498	DEFINITION	Mesocricetus auratus telomerase catalytic subunit mRNA, complete	
2401	TTCTTCTCTGCACTTCTGCTGCTCAGAGTGTCTGAAGATTTGTTGATGAGTGTG	2460	ACCESSION	AF149012	
2499	TGCGAGGGCATCCCGAGGCTCCAGCTATCCACCTGCTCTGAGTCTGTTTCGGA	2558	VERSION	AF149012.1	
2461	TGCGAGGGCATCCCGAGGCTCCAGCTATCCACCTGCTCTGAGTCTGTTTCGGA	2520	KEYWORDS	GI:6572639	
2559	GACATGGAGAAACAGCTGTTTCTGCTGAGGTGACGGGATGGGTTGCTTTACGTTTGT	2618	SOURCE	Mesocricetus auratus (golden Syrian hamster)	
			ORGANISM	Mesocricetus auratus	
				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;	
				Mesocricetus.	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
EDLINE  
PUBMED  
21240330  
2 (bases 1 to 4170)  
Guo, W., Okamoto, M., Lee, Y.M., Baluda, M.A. and Park, N.H.  
Enhanced activity of cloned hamster TERT gene promoter in  
transformed cells  
Biophys. Acta 1517 (3), 398-409 (2001)

2 (bases 1 to 4170)  
Guo, W., Okamoto, M. and Park, N.H.  
Direct Submission  
Submitted (06-MAY-1999) Dentistry, University of California, Los  
Angeles, 10833 Le Conte Ave., Los Angeles, CA 90095, USA  
Location/Qualifiers  
1..4170

source  
/organism="Mesocricetus auratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10036"  
/notes="common: Golden Syrian hamster"  
28..3414  
/notes="hamTERT"  
/codon\_start=1  
/product="telomerase catalytic subunit"  
/protein\_id="AA17334.1"  
/db\_xref="GI:557240"  
/translation="MPAPRCNVAALLRSQYRQVWPLATFVRLGPGRQLVQGGDP  
KVFRLVACLVCPWDSQPPADLSHOVSSLELVARVQRLCBERGNVLTGFA  
LVNAGGGPMFTTYSYSLPNSVTSRVSAGMILLNVRGDDLLVYLARCALYL  
LPPSCAYQCGSPFTYQICATETWPSVSYIYRTPVGRNTHLGSFVRNSSHQ  
ANKPPLPREAKRSLSTNRSVPSKARCDLAPLEKGPYQAVTPSPDXTWVNP  
AKSHAVPSRTTKEDLSGKAPGLSRSGVCYKHPSTLSQSLQNAFOLRYPTE  
TKRFLYRSGGERLNPSPFLNNLQPSLTGARRLVEILPLGMRPTSGPLCGRRLSK  
RYWQPLRQQLLVNARCPVRLRLSHRFRTPAHOAGALNTTSPORLNLRLHS  
SPQVYVFLQACVGLQVPLGLWGRHNRQFEKVRFEISLQYKXLSQELTWMKY  
QDCWLRSSPNNCPVAEHTREILAEFLMLDAYVVELLSFTYETTFQKRL  
LEFYKMRRLQSLGVRRHRLVRLQELSOEVRQEAWPAMPICRURFIPKESGL  
RTFVNVNLTGAPDFKQAOHQFTQCLTFLSVNLKELTHTLNLGASVGLNDLYRT  
WTFVLRVLTADPFRFVADVTGAYDAIPQKLVEVTIANMRHPDINSYCHQYAV  
VDRDQCIHKSFRRQVSTLSLQPHMGQFLKLDSDTSALRNSVVIQSLSLNEAS  
SFLDFPFLFVANSVVKYGGCYVQCGIPOGSLSTLLCLSCFDMENKLPFAEQOD  
GLLRDDDFLLVTLPHLVQABFRLADLVGTPYEGCMINIKQTVNFPVDAGTLQTA  
HPLRPHCLFPMQGLLQTLQLEVLCDYVARTSIKASLTFTQPKAGNNRQKLLA  
VLRLKCHSLFLDLQMSLQTVINVKIFLLOAYRFHACALQLPDQHVKNPAPFLS  
LSINIASCCYSILKVNAGMTLKGASGFFPEAPRWUCCITQAFLLKLAGHSVITKCL  
LGPLRTAQKQCRKPRATMAILETAADPALSTDQFILD"

ASE COUNT 927 a 1177 c 1126 g 940 t  
RIGIN

Query Match 73.8%; Score 2578.6; DB 10; Length 4170;  
Best Local Similarity 85.4%; Pred. No. 0;  
Matches 2941, Conservative 0; Mismatches 479; Indels 25; Gaps 5;  
12 GGGAGGCCATCCGCGCTTGACACATGACCCGCGCTCTCGTTCCCGCGGTGGCG 71  
1 GGGAGGCCCGCGCGATCTGACCGCGATGCCCGCGCGCGCGCGCGCGCGCGCG 60  
72 TCTCTGTCGGCAGCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTG 131  
61 GCTCTGTCGGCAGTCAATACCGTCAAGTGTGGCGCTGGCAACCTTTGTGGCGCGCTG 120  
132 GGGCCGAGGCGAGCGGTGTGGCAACCGGAGCCCGGAGATCAACGACCTTTGTT 191  
121 GGACTGAGGCGAGGAGCTGTGACACCGGAGCCCAAGGCTTCCTCCGACGTTGGTG 180  
192 GCCCAATGCTAGTGTGATGATGATGCGGCTCACAGCTCCACCTGCGGACCTTTCTTTC 251  
181 GCCCGGTGCTAGTGTGTGGCTGGGACTCACACCTCCACCTGCTGACCTTTCTTTC 240  
252 CACGAGGTGTCTCTTGAAGAGCTGTGGCGAGGTTGTCAGAGACTCTCGAGGCG 311  
241 CACGAGGTGTCTCTTGAAGAGCTGTGGCGAGGTTGTCAGAGACTCTCGAGGCG 300  
312 AACGAGAGAAAGTGTGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGGCGGCGCT 371

301 GCGGAGAGAAAGTGTGATCTTTTGGCTTGGCTGCTTAAAGGAGCCCAAGGCGGTCT 360  
372 CCCATGGCTTCACTAGTGTGCTGCTAGTCTTCCCAACACTGTTATTAGAGACCTG 431  
361 CCATGACATTCACACAGGCTGCGAGCTACTGCCCCAACTCGGTGACTGACTCTG 420  
432 COTGTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491  
421 CCGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
492 CTGCTGGGACACTGTGCTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551  
481 CTGCTGGGCGCTGTGCTGCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
552 GGGTCTCCCTGCTACCAATTTGTCACCAAGATATCTGGGCTCTGTGCTGCTGCTGCT 611  
541 GGTCTACCTCTGTACCAATTTGTCACCAAGATATCTGGGCTCTGTGCTGCTGCTGCTG 600  
612 TACAGGCGCCACCGGCTGGGAGGATTTTCACTAAGCTTAGTGTCTTACACAGATC 671  
601 TACAGGCGCCACAGACCGCTGGGAGGATTTTCACTAAGCTTAGTGTCTTACACAGATC 660  
672 AAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731  
661 AGAAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
732 AGGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791  
721 CGGAGTCTAAGCATCAACCAATAGAGTGTGCTTCACTTAAAGAGCCAGGTGCGATCT 780  
792 GTCCGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 851  
781 GCGCGAGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
852 TGGGTGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902  
841 TGGGTGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
903 TGTGCTTCTTAAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962  
901 TGTGCTTCTTAAAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960  
963 AAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022  
961 AAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
1023 TTTATTGAGACGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1082  
1021 TATCTGAGACGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1080  
1083 TCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1142  
1081 TGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1140  
1143 ATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1202  
1141 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
1203 CGTCACTACTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1262  
1201 AAGCGCTACTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
1263 CAATATGTCAGATCTCTCAGGTCACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1322  
1261 CGGTATGTGCGGCTCTCTCAGGTCACTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320  
1323 GATGCTGCTGAA---CACCAGGCGCCAGCTCATGATTTGCTCGGCTCTGACAGCAGT 1379  
1321 GGTGCTTGAACACACACAGGCTCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1380  
1380 CCGTGGCAGGTATATGTTTCTTGGGCTGCTGCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCT 1439  
1381 CCGTGGCAGGTATATGTTTCTTGGGCTGCTGCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCT 1440

1440 TGGGGTACAGGCAAAATGAGCGCGCTTCTTTAAGAACTTAAAGAAAGTTCACTCTCGTTG 1499  
1441 TGGGGTTCGCGCAAAACAGGAGCGTCTTTAAGAAAGTGAAGCGTTCACTCTCTTG 1500  
1500 GGGAAATACGGCAAGCTATCACTGCAAGGACTGATGTGGAGATGAAGTAGAGATTGC 1559  
1501 GGGAAATATGAACAAGTCTGCTGCAAGAGCTGACGTGGAGATGAAGTTCAGAGACTGC 1560  
1560 CACTGGCTCCGAGCAGCGCCCGGGAAGAACCGTGTCCCGCTGCGAGAGCAACCGTCTGAGG 1619  
1561 AGTGGCTTCGAGCAGCGCCCGGGAAGAACCGTGTCCCGCTGCGAGAGCAACCGCAGG 1620  
1620 GAGAGATCTGGCTTACGTTCTGTTCTGGTGATGACACATACGTGGTACAGCTGCTT 1679  
1621 GAAAGATCTGGCTTGTCTTCTGTTCTGGTGATGACGCGTACGTGGTAGAGCTGCTT 1680  
1680 AGTCAATCTTTTATACACAGAGAGCACTTCCAGAAAGAACAGGCTCTTCTTCTACCGT 1739  
1681 CGGTCAATCTTTTACGTACAGAGACCACTTCCAGAAAGAACCGGCTCTTCTTCTACCGA 1740  
1740 AAGAGTGTGGAGCAAGCTGCAAGCAATCGAATCGAGCAACACCTTGAGAGAGTGGG 1799  
1741 AAGAGCATGTGGAGAGGCTGCAAGCAATCGAATCGAGCAACACCTTGAGAGAGTGGG 1800  
1800 CTACGGGAGCTGTCAAGAGAGGCTCAGGCATCACCAGGACACCTGGCTAGCCATGCC 1859  
1801 CTACAGAACTGTCTCAAGAGAGTCAAGGAGGCTGAGGAGGCTGGCAGGCCATGCC 1860  
1860 ATCTGAGACTCGGCTTCATCCCAAGCCCAAGCGCTCGCGCCCATTTGTGAACATGAGT 1919  
1861 ATCTGAGACTCGGCTTCATCCCAAGCCCAAGCGCTCGCGCCCATTTGTGAACATGAGT 1920  
1920 TATAGCATGGTACAGAGCTTTGGGAGAGAGAGAGGCGCCAGCAATTTTCAACCGCT 1979  
1921 TAT---ATGGGACACAGAGCTTTGACAAAGGAGAGGCTCAGCAATTTTCAACCGCT 1977  
1980 CTCAGACTCTCTTCAGCATGTCTCAACTATGAGGGGCAAAACATCTCTCACTTTATGGG 2039  
1978 CTCAGACTCTCTTCAGCGTCTCAACTATGACTGACAAACATCTCACTTTCTGGG 2037  
2040 TCTTCTGTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2089  
2038 GCATCTGTAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2097  
2100 CFTGCTCTGACAGACACCCAGGATGATTTGTTAAGGAGAGATGATGATGATGATGATGATG 2159  
2098 CGCATCTGACAGACACCCAGGATGATTTGTTAAGGAGAGATGATGATGATGATGATGATGATG 2157  
2160 GATGCCATCCCCAGGTAAGTGTGTGAGGTTGTTGCCAATATGATGATGATGATGATGATGATG 2219  
2158 GATGCCATCCCCAGGTAAGTGTGTGAGGTTGTTGCCAATATGATGATGATGATGATGATGATG 2217  
2220 AGCAGTCTGATTCGCGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2279  
2218 AACTCTGATGATTCACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2277  
2280 AAGTCTTTAGGAGACAGGTACACCTCTCTGACCTCCAGCCATACATGGGCGAGTTC 2339  
2278 AAGTCTTTAGGAGACAGGTACACCTCTCTGACCTCCAGCCATACATGGGCGAGTTC 2337  
2340 CTTAAGCATCTGAGGATTCAGATGCGAGTGCATGAGGAACTCGGTGATGATGATGATGATGATG 2399  
2338 TTAAGCATCTTCAGGACTTCAGACCAAGTGGCTGAGGAACTCGGTGATGATGATGATGATGATG 2397  
2400 AGCATCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2459  
2398 AGCTTATCTGAGAGGCGAGGAGGAGGCTGTTTCTGATCTTCTCTCTCTCTCTCTCTCTCTCT 2457  
2460 CAGAGTGTGTAAGATTTGGTGAAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2519  
2458 AACAGTGTGTAAGATTCGGTGGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATG 2517

QY 2520 TCCAGCCTATCCACCTGCTCTGCACTCTGTGTTTCGGAGACATGGAGAACAGCTGTTT 2579  
Db 2518 TCCAGCCTGTCACCTGCTCTGCACTCTGTGTTTCGGAGACATGGAGAACAGCTGTTT 2577  
QY 2580 GCTCAGGTGAGCGGATGGTTGCTTTTACGTTTGTGATGACTTTCTGTTGTGAGC 2639  
Db 2578 GCTCAGGTGAGCGGATGGTTGCTTTTACGTTTGTGATGACTTTCTGTTGTGAGC 2637  
QY 2640 CCTCAGTGGACCAAAACCTTCTCAGACCCCTGTCAGACCCCTGTCAGTGTGATGAT 2699  
Db 2638 CCTCAGTGGTCCAGCGGAGGCTTCTCAGCGGCTCTGTCGTCGTCATCCCTGAGTAC 2697  
QY 2700 GGGTGCATGATAAATCTGAGAGACAGTGGTGAATCTCCCTGTGAGCTGTGATCCCTG 2759  
Db 2698 GGGTGCATGATAAATCTGAGAGACAGTGGTGAATCTCCCTGTGAGCTGTGATCCCTG 2757  
QY 2760 GGTGCTCAGCTCCATACAGCTGCTGCTCACTGCTGTTTCCCTGTGAGCTGTGATCCCTG 2819  
Db 2758 GATGGCAGCTCCACACAGCTGCTGCTCACTGCTGTTTCCCTGTGAGCTGTGATCCCTG 2817  
QY 2820 CTGACACTCAGACTTGGAGGTGTTCTGACTACTCAGGTATATGCCAGACCTCAAT 2879  
Db 2818 CTGACACTCAGACTTGGAGGTGTTCTGACTACTCAGGTATATGCCAGACCTCAAT 2877  
QY 2880 AAGACGAGCTTCCAGAGTGTCTTCAAGCTGGGAAGACCATGCGGAACAAGCTC 2939  
Db 2878 AAGCCAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTTCCAGAGCTC 2937  
QY 2940 CTGTCGCTTGGGTTGAAGTGTGAGGTCTCAGGTCTATTTCTAGACTTTCAGGTGAGACGCTC 2999  
Db 2938 TTAGCTGTTTGGGTTGAAGTGTGAGGTCTCAGGTCTTCTGACTTTCAGGTGAGATGATGCTT 2997  
QY 3000 CAGACAGTCTGCATCAATATATACAGATCTTCTGCTTTCAGGCTTACAGGTTTCCATGCA 3059  
Db 2998 CAGACAGTCTGCATCAATATATACAGATCTTCTGCTTTCAGGCTTACAGGTTTCCATGCG 3057  
QY 3060 TGTGTGATGACTTCCCTTTGAGAGCTGTTTGAAGAACCTCATTCTTCTGAGC 3119  
Db 3058 TGTGCTTTCAGCTTCCCTTTGAGAACCATGTTTGAAGAACCTTCTTCTGAGC 3117  
QY 3120 ATCATCTCCAGCAAGCATCTGCTGCTATGCTATCTGAAAGTCAAGATCCAGGAATG 3179  
Db 3118 ATCATCTCCAGCATGATCTGCTGCTTCTGCTTCTGCTTCAAGGTCAAGATGAGGAATG 3177  
QY 3180 ACATTAAG-----GCTCTGGCTCTTCTTCTTCTGAGAGCCGCAATGCTCTG 3230  
Db 3178 ACATTAAGGCAAGGCTGCTCTGCTCTTCTTCTTCTGAGCTGCACTGCTCTG 3237  
QY 3231 TACCAGGCTTCTGCTCAAGCTGGCTGCTCATCTGCTCATCTACAAATGCTCTCTGGGA 3290  
Db 3238 TACCAGGCTTCTGCTCAAGCTGGCTGCTCATCTGCTCATCTACAAATGCTCTCTGGGA 3297  
QY 3291 CTTCTGAGGACAGGCGGCAAACTGCTGCTGGGAAGCTCCAGAGGCGCAATGACCATC 3350  
Db 3298 CTTCTGAGGACAGGCGGCAAACTGCTGCTGGGAAGCTCCAGAGGCGCAATGAGGCGCATC 3357  
QY 3351 CTTAAGGTGAGTGAAGGCGGCTTGAAGCAGAGCTTTCAGAGGCTTTCAGAGGCTTTCAGGCT 3410  
Db 3358 CTTGAGTCTGCTGAGGCGGCTTGAAGCAGAGCTTTCAGAGGCTTTCAGAGGCTTTCAGGCT 3417  
QY 3411 TGTCTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3435  
Db 3418 -CTCCCTTTCAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3441

## RESULT 4

AF247818  
LOCUS Rattus norvegicus telomerase catalytic subunit mRNA, partial cds.  
DEFINITION  
ACCESSION AF247818  
VERSION AF247818.1 GI:7407652  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)

ORGANISM	Rattus norvegicus	Db	421	CGTCTCTGGACACACCCAGGATGTACTTTTAAAGCAGATGTACCCGGGGCTAT	480
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
TITLE	Rattus.	Qy	2160	GATGCCATCCCCAGGGTAAAGCTGGTGGAGGTTGTGCCAATATGATCAGGCACCTCGGAG	2219
JOURNAL	1 (bases 1 to 1758)	Db	481	GATGCCATCCCCAGGGTAAAGCTGGTGGAGGTTGTGCCAATATGATCAGGCACCTCGGAG	540
REFERENCE	Wong, S., Gao, S., Xu, X. and Yu, H.	Qy	2220	AGCAGTACTGTATCCCGCCAGTATGAGTGGTGGAGAGATAGCCAGGCCAAGTCCAC	2279
AUTHORS	2 (bases 1 to 1758)	Db	541	AGCAGTACTGTATCCCGCCAGTATGAGTGGTGGAGAGATAGCCAGGCCAAGTCCAC	600
JOURNAL	Wong, S., Gao, S., Xu, X. and Yu, H.	Qy	2280	AAGTCTTTAGGACAGAGTCCACACCCCTCTCTGACCTCCAGCCATACATGGGCCAGTTC	2339
TITLE	Submitted (21-MAR-2000) IMRE, National University of Singapore, 10	Db	601	AAGTCTTTAGGACAGAGTCCACACCCCTCTCTGACCTCCAGCCATACATGGGCCAGTTC	660
ATRRS	Medical Drive 117597, Singapore	Qy	2340	CTTAAGCATCTGCGAGGATTCAGATGCGAGTGCATGAGGAATCCCGTGTCTATCGAGCAG	2399
source	Location/Qualifiers	Db	661	CTTAAGCATCTGCGAGGATTCAGATGCGAGTGCATGAGGAATCCCGTGTCTATCGAGCAG	720
CDS	1..1758	Qy	2400	AGCATCTCTATGAATGAGAGCAGCAGCAGCAGTGTTCGACTTCTTCTGCACTTCTCTGCGT	2459
	/organism="Rattus norvegicus"	Db	721	AGCATCTCTATGAATGAGAGCAGCAGCAGCAGTGTTCGACTTCTTCTGCACTTCTCTGCGT	780
	/mol_type="mRNA"	Qy	2460	CACAGTGTCTGTAAGATTGTTGACAGTGTCTATACGAGTGCAGGGCATCCCCCGAGGC	2519
	/db_xref="taxon:10116"	Db	781	CACAGTGTCTGTAAGATTGTTGACAGTGTCTATACGAGTGCAGGGCATCCCCCGAGGC	840
	<1..1728	Qy	2520	TCCAGCTTATCCACCTCTCTGAGTGTCTGTTTCCGAGACATGGAACAACAGCTGTTT	2579
	/note="rTER"	Db	841	TCCAGCTTATCCACCTCTCTGAGTGTCTGTTTCCGAGACATGGAACAACAGCTGTTT	900
	/codon_start=1	Qy	2580	GCTGAGGTGCAGCGGATGGTGTCTTTTACGTTTCTTGTGATGACTTCTTGTGGTGAAG	2639
	/product="telomerase catalytic subunit"	Db	901	GCTGAGGTGCAGCGGATGGTGTCTTTTACGTTTCTTGTGATGACTTCTTGTGGTGAAG	960
	/protein_id="AAF62177.1"	Qy	2640	CCTCATCTGGACCAAGCAAAAACCTTCTCAGACCTGCTGCTCCATGGCGTTCCTGAGTAT	2699
	/db_xref="GI:7407653"	Db	961	CCTCATCTGGACCAAGCAAAAACCTTCTCAGACCTGCTGCTCCATGGCGTTCCTGAGTAT	1020
	/translation="RSFFVITESTQKURLFFYRKSVMKLSIGVROHLERVLREL	Qy	2700	GGTGTGATGATAAACTTGCAGAACAGTGGTGAACCTTCCCTGTGGAGCCTGGTACCCCTG	2759
	SQEVRRHQDTWLMPIELRIPXNGRLPIVMSYSMTGRALGRKQAHPTORLK	Db	1021	GGTGTGATGATAAACTTGCAGAACAGTGGTGAACCTTCCCTGTGGAGCCTGGTACCCCTG	1080
	TLFSMLNERTKHPHLMGSSVGMNDIIRTWRAFVLRVLRADQTPRMFYKADVTGX	Qy	2760	GGTGTGATGATCCATACAGCTGCTGCTCAGTGTCTTCCCTGCTGGTGGCTTGGCTG	2819
	DAIPQKJVEVNMIRHSESTCIRQAVVRDSQGVHKSFRQVYTTLSDLQPMG	Db	1081	GGTGTGATGATCCATACAGCTGCTGCTCAGTGTCTTCCCTGCTGGTGGCTTGGCTG	1140
	QFLKHLQDSALRNSVIEOSISMNESSLDFLFLHRHVSIVRCRYTCQG	Qy	2820	CTGGACACTCAGACTTTCGAGGTGTCTGTGACTACTCAGGTATGCCCAGACCTCAATT	2879
	IPGSLHSLTLLSCFGDMENKLFVAVQDGLLRFVDFDLVFLPHLDORAKTFLSV	Db	1141	CTGGACACTCAGACTTTCGAGGTGTCTGTGACTACTCAGGTATGCCCAGACCTCAATT	1200
	HGVEYGMINKLQTVNPFVPGTILGAAQYLPANCLFPWCGLLDTOTLEVCDDY	Qy	2880	AAGACAGAGCTCACCCTTCAGAGTGTCTTAAAGCTGGGAAGACCATGCGGAACAAGCTC	2939
	SGVAQNSIKTSLTQSFVKGAGTMKLLSVLRKCHGLFDLDVNSLQTVNINIKI	Db	1201	AAGACAGAGCTCACCCTTCAGAGTGTCTTAAAGCTGGGAAGACCATGCGGAACAAGCTC	2999
	FLLOARFRACVIOIPFDORVKNLTFPLGIISSOACCVAILKVPNGMTLKASGFF	Qy	1261	CTGTGGTCTTCCGGTTGAGTGTCAAGTGTCTATTTCTAGACTTTCAGAGGTGAACAGCTC	1320
	PPERAHWLCYQFLKLAHSAVYIKLLGLPLNTAQKLCRLKPLPENTWILKAADPAL	Db	3000	CAGACAGTCTGCATCAATATATACAAAGATCTTCTGTCTAGGCTTACAGGCTTCCATGCA	3059
	STDFQRIILD"	Qy	1321	CAGACAGTCTGCATCAATATATACAAAGATCTTCTGTCTAGGCTTACAGGCTTCCATGCA	1380
misc_feature	<1..87	Qy	3060	TGTGTGATTCAGCTTCCCTTTGACCCAGCGTGTAGGAGAACCTCAGATCTTCTTCTGGC	3119
SE COUNT	398 a 486 c 447 g 427 t	Db	1381	TGTGTGATTCAGCTTCCCTTTGACCCAGCGTGTAGGAGAACCTCAGATCTTCTTCTGGC	1440
IGIN	/notes="Region: motif T"	Qy	3120	ATCATCTCCAGCAAGCATCTCTGTCTATGCTATCTCTGAAAGGTCAAGAAATCCAGGAATG	3179
Query Match	50.2%; Score 1756; DB 10; Length 1758;	Db	1441	ATCATCTCCAGCAAGCATCTCTGTCTATGCTATCTCTGAAAGGTCAAGAAATCCAGGAATG	1500
Best local Similarity	100.0%; Pred. No. 0;	Qy	3180	ACACTAAGGCTCTGGCTCTCTTCTCTGAGCCGACATTTGGCTCTGCTACAGGCC	3239
Matches 1756; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
1680	AGGTCACTCTTTTACATCAGAGAGCACATCCAGAGAACAGGCTCTTCTTACCGT				
1	AGGTCACTCTTTTACATCAGAGAGCACATCCAGAGAACAGGCTCTTCTTACCGT				
1740	AAGAGTGTGGAGCAAGCTGACAGACATTTGAGTCAAGCAACACCTTGGAGAGTGGG				
61	AAGAGTGTGGAGCAAGCTGACAGACATTTGAGTCAAGCAACACCTTGGAGAGTGGG				
1800	CTACGGAGCTGTACAGAGAGGAGTCAAGCAACACCTTGGAGTACCCATGCCC				
121	CTACGGAGCTGTACAGAGAGGAGTCAAGCAACACCTTGGAGTACCCATGCCC				
1860	ATCTGACAGTGGCTTATCCCAAGCCCAACAGGCTGCGGCCCATTTGTGAACATGAGT				
181	ATCTGACAGTGGCTTATCCCAAGCCCAACAGGCTGCGGCCCATTTGTGAACATGAGT				
1920	TATAGCATGGTACCAAGCTTTGGGCAAGAGCAAGCAAGCCAGCATTTCAACCCAGGCT				
241	TATAGCATGGTACCAAGCTTTGGGCAAGAGAGCAAGCCAGCATTTCAACCCAGGCT				
1980	CTCAAGACTCTCTTTCAGCATGCTCAATCATATGAGCGGACAAAACATCTTATGGG				
301	CTCAAGACTCTCTTTCAGCATGCTCAATCATATGAGCGGACAAAACATCTTATGGG				
2040	TCTTCTGTATGGGTATGATGATCATCAGACACTGGCGGCCCTTTGTGCTGGTGTG				
361	TCTTCTGTATGGGTATGATGATCATCAGACACTGGCGGCCCTTTGTGCTGGTGTG				
2100	CGTGTCTGGACACACCCAGGATGTACTTTTAAAGCAGATGTGACCCGGGCTAT				



b 1501 ACACCTAAAGCGCTGCTGGCTCTCTTCTCTCTGAGCGGCACATGCGCTCTGCTACAGGCC 1560  
y 3240 TTCTCTGCTCAAGTGGCTGCTCATCTCTGCTCATCTACAAATGCTCTCTGGGACCTCTGAGG 3299  
b 1561 TTCTCTGCTCAAGTGGCTGCTCATCTCTGCTCATCTACAAATGCTCTCTGGGACCTCTGAGG 1620  
y 3300 ACAGCCCAAACTGCTGTCGCGGAGCTCCAGAGGCGACAATGACCATCTTAAAGCT 3359  
b 1621 ACAGCCCAAACTGCTGTCGCGGAGCTCCAGAGGCGACAATGACCATCTTAAAGCT 1680  
y 3360 GAGCTGACCCAGCCTTAAAGCAGACATCTTACAGACATCTTGGACTAACCTGCTCTCTT 3419  
b 1681 GAGCTGACCCAGCCTTAAAGCAGACATCTTACAGACATCTTGGACTAACCTGCTCTCTT 1740  
y 3420 CCGCTAGTAGACATG 3435  
b 1741 CCGCTAGTAGACATG 1756  
  
RESULT 5  
LOCUS AF043739  
DEFINITION Synthetic construct human telomerase catalytic subunit (hTERT/hEST2) mRNA and 3' Influenza A hemagglutinin tag, complete sequence.  
ACCESSION AF043739  
VERSION AF043739.1 GI:3023054  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D., Haber,D.A. and Weinberg,R.A.  
TITLE Telomerase activity is restored in human cells by ectopic expression of hTERT (hEST2), the catalytic subunit of telomerase  
JOURNAL Oncogene 16 (9), 1217-1222 (1998)  
MEDLINE 98187713  
PUBMED 9528864  
REFERENCE 2 (bases 1 to 8960)  
AUTHORS Counter,C.M., Meyerson,M., Eaton,E.N., Ellisen,L.W., Caddle,S.D., Haber,D.A. and Weinberg,R.A.  
TITLE Direct Submission  
JOURNAL Submitted (20-JAN-1998) Whitehead Institute, 9 Cambridge Center, Cambridge, MA 02142, USA  
FEATURES  
source  
1..8960  
Location/Qualifiers  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/focus  
order(1..1101,4606..8960)  
source  
/organism="Cloning vector pCI-neo"  
/db\_xref="taxon:45772"  
/notes="mammalian expression vector; see GenBank Accession Number U47120"  
1102..4542  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/map="5p15.33"  
/notes="telomerase catalytic subunit hTERT/hEST2 mRNA"  
4543..4605  
source  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/notes="engineered based on the epitope for the anti-influenza virus hemagglutinin monoclonal antibody 12CA5; similar to part of the Influenza A hemagglutinin gene sequence in GenBank Accession Number M29257"  
1102..4542  
gene  
/gene="hTERT/hEST2"  
1147..4605  
CDS  
/note="telomerase/reverse transcriptase"  
/codon\_start=1

/transl\_table=11  
/product="telomerase catalytic subunit"  
/protein\_id="AAC40212.1"  
/db\_xref="GI:3023055"  
/translation="MPAPPCRAVRSLLRSHYREVLPLATFVRLRGPQGRWLVRQGDPA  
AFALYAOQLCVCPWDAPPPAPSPROVSCLELVARVLOBLRERGAKNVLAFGFA  
LLDARGGFPEAFVTSVSYLPTNTDALLGSGAWGLLLRRVDDDVVHLLARCALFV  
LVAPRQYVCGGPLYQLGAATQARPPPHASGRRLRGCERAWNHVSRRGFCVSPGA  
PGARRGASRSRLPLPKPRRGAAPERTPVQGSWAHPGRTGRPSDRGVFVSPFA  
RPAEATSLRGLSGTRHSHPSVGRHAGFPSTPRPFMDTTCPPVVAETKHFLYS  
SGDEQLRPFLSLRPSLTGARRLVETIFLGRPMWPTPRRLPRLPORYWQMRPL  
FLELHSHQACPYGVLLKTHCPRLAAVTPAAGVCAREKPOGSVAAPBEEDTDPRLVQ  
LLRHSHPWQVYGRACRELVPGLGSHNERELRNTKKFISLGHAKLSLQEL  
TWKSVRGCAWLRSPGVCPVPAERLREILAKFLHMLSVYVELLARSFYVETET  
TFQNRLLFYRMSVSKLQSIGIRQHLRVQLRSEAEVRQKREARPAULLTRLRFI  
PKPDLRPIVNMDDYVVGARTFRREKABRLTSRKALFSVLYNERARPOLLLGASVLG  
LDDIHRARTFVLRRAQDPPELYFVKVDVTGAYDTIPQDLRTVEIASIIKQNTYC  
VRYAVVQKAAHGRVAKFSVSTLTDQVYMRQFVAHQETSPLEDAVVEQSLS  
NEASSGLFDVFLREMHAVIRGKSVQCGIPQGSILSTLLSCYGDGMENKLFAG  
TRDGLLRVDDDELLVTHAKTFLRVRGVEYGVNLRKTVNVPVEDEAL  
GCTAFVQMPAGLFPWCGILLDTRELYQSDYSYARTSRASLTENRFGKAGNNR  
KLFGVRLKCHSLFDLQVNSLQVCTNIIKILLQYRFHACVLQPLPHQQWKNPT  
FLEIVSDTASLCVSLIKNAGNSGAGAGLPSEAVQWLC HQALFKLTKLTHRTVT  
YVPLGLSLRTAQTLQSLRKLFGTLLTALEAAANPALPSPDKTILDPYDVPDYAGPYD  
VPDYAG"

BASE COUNT 1886 a 2551 c 2484 g 2038 t 1 others

ORIGIN

Query Match 45.9%; Score 1603.6; DB 12; Length 8960;  
Best Local Similarity 68.7%; Pred. No. 0;  
Matches 2389; Conservative 0; Mismatches 979; Indels 108; Gaps 9;

QY 23 CCGCGCTTGAGCAATGACCCCGCTCTCTGTTGCCCGCGGTGGCTCTCTGCTGCG 82  
DB 1131 CCGCGCACCCCGCGATGCCGCGCTCCCGCTGCCGAGCGCTGCTGCTGCG 1190  
QY 83 CAGCGATACCGGAGGTGGCGCTGGCAACCTTTGTGGCGGCTGGGGCCCGAGGG 142  
DB 1191 CAGCCACTACCGGAGGTGGCGCTGGCGCTGGCGCTGGCGGCGCCCGCCAGGG 1250  
QY 143 CAGGGGCTTGTGCAACCGCGGACCGAGATCTACCGCACTTTGTTGCCCAATGSCCT 202  
DB 1251 CTGGCGCTTGTGACGCGCGGACCGCGGCTTCCGCGGCTGGTGGCCAGTGCCT 1310  
QY 203 AGTGTGCATGCACTGGGCTCAGAGCTCAGAGCTTCCACTGCCACCTTTCTTCCACAGGTCT 262  
DB 1311 GGTGTGCTGCTTGGGACGACGCGCGCCCGCCCGCTTCTCGCGAGGTCTC 1370  
QY 263 ATCCCTGAAGAGCTGTGGCCAGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322  
DB 1371 CTGGCTGAAGAGCTGTGGCGCGGAGTGTGACAGGCTGTGCGAGCGCGCGCGAGAA 1430  
QY 323 CGTGTGCTTTTGTGCTTGTAGCTTGTAAACAGGCGCAGAGCGGGGCTTCCCATGGCCTT 382  
DB 1431 CGTGTGCTTGTGCTTGTGCTGCTGCGCTGCGAGCGGCGCGCGGCGCGCTT 1490  
QY 383 CACTAGTGTGCTGTAGTACTTGTGCGCAACACTGTTATTGAGACCTGCTGTGCTCAGTG 442  
DB 1491 CACCAACGCTGTGCGAGCTTGTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 1550  
QY 443 TGCATGATGTCTACTTGTGAGCGGAGTGGGCGACGACTGCTGCTGTCTACCTGTGGCACA 502  
DB 1551 GCGGTGGGCGCTGCTGTTGCGCGGCTGGGCGACGAGTGTGCTGTGCTGCTGGCAGC 1610  
QY 503 CTGTGCTTTTATCTTCTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCCCT 562  
DB 1611 CTGCGGCTTCTTGTGCTGCTGCTTCCAGCTGCGCTTCCAGGTGTGCGGCGCGCGCT 1670  
QY 563 GTACCAAAATTTGTGCGCACACGATATCTGCGCGCTCTGTGCTCGCTAGTTACAGGCCAC 622  
DB 1671 GTACCAGCTGCGCGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1730  
QY 623 CCGACCGCTGGCGAGGAATTTCACTAAGCTTGTGCTTCTTACACAGATCAAGAGCAGTAG 682

1731 GGCTGCGATGC-----GAACGGGCTGGAACCATAG 1763  
683 TCGCAGAGAGACCGAAACCCCTGGGCTTGGCATCTCGAGGTACAAAGAGGCACTCTGAG 742  
1764 CGTACGGAGGCGCGGGTCCCTTGGGCTTGGCAGCCCGGGTGGAGAGGCGCGGGGG 1823  
743 TCTACAGTACAGAGTGGCTTCACTAAGAGGCGAGATCTATCTCTGCTCCGAGAGT 802  
1824 CAGTGCACAGCAAGTCTGCGGTTGCGCAAGAGGCGCGAGGCTGGCGCTGCCCTTGAGCC 1883  
803 GGAGGAGGACCCACAGCAGGTGCTTACCAACCCATCAGGCAAAATCATGGGTGCCAAG 862  
1884 GGAGCGGACCGCTGGTGGCAGGGGCTCTGGGCGCCACCGGCGAGGACGCGTGGACGAG 1943  
863 TCTGCTGGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTTGTTCTTTAA 913  
1944 TGACCGTGGTTCTGTGGTGGTCACTGCGGACCGCGGAGAGGCAACCTCTTTGGA 2003  
914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGSGTGGGTGCTGTAAACACAAAGCCGAG 970  
2004 GGGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGGCGCCAGCACCGCGGG 2063  
971 CTCCACAT---CTGCTGCTCACACCCCGCAAAATGCTTTAGCTCAGGCCATTTAT 1027  
2064 CCCCCTCATCATCGGCGCCACCATCTCTCTGGACACGCTTGTCTCCCGGTGTAGCG 2123  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGGCTCTAAACCCCTCAT 1087  
2124 CGAGACCAAGCACTTCTCTACTCTCTCAGCGCA---CAAGAGAGCAGCTGCGGCGCTCTT 2180  
1088 CTTACTCAGCAACCTCCAGCCTTAACCTTGACTGGGGCAGGAGACTGGTGGAGATCATCTT 1147  
2181 CTTACTCAGCTCTCTGAGGCCACGCTGCTGCGCTCGAGGCTCTGTGGAGACATCTT 2240  
1148 TCTGGGCTCAGGCTTAGGACATCAGAGCACCTCTCGAGACACACCTCTATGGGTGG 1207  
2241 TCTGGGTTCCAGGCGCTGGATGCGCAGGGACTCCCGCGAGGTTGCGCCCGCTGCGCCAGCG 2300  
1208 ATACTGGCAGATGGGCGCTTGTTCACAGCTGCTGTGTGAACCAATGAGAGTGGCCAAATA 1267  
2301 CTACTGGCAATGGGCGCTTGTTCAGAGCTGCTTGGAGTCTTGGGAACCAAGCGGAGTSCCCTA 2360  
1268 TGTACAGCTCTCTAGGTCACATTCAGAGTTTCAAGAGGTTTCAAGCAACCAAGGTGACAGATGC 1327  
2361 CGGGTGTCTCTAAGACGCACTCCCGCTGCGAGCTGCGGTCAACCCAGCAGCGGGTGT 2420  
1328 CT-----TGAAACCAAGCCC 1342  
2421 CTGTGCGCGGAGAACCCAGGCTCTGTGGGCGCCCGCGAGGAGGAGACAGAGCCC 2480  
1343 ACCGACCTCATGATTTGCTCCGCTGACAGAGTCTCTGCGAGGTATATGGTTTCT 1402  
2481 CGTCCGCTGGTGCAGCTGCTCCGCGACACAGAGCCCTCTGCGAGGTGTACGGCTTCGT 2540  
1403 TGGGCTGTCTCTCAGAGTGTGTCTGTAGTCTCTGGGTACAGGCAATAGAGCG 1462  
2541 GCGGCGCTTCTGCGCGCTGGTGGGCTCTGCGGCTCTCGAGGCAACAGBAACG 2600  
1463 CGCTTTCTTTAAGAACTTTAAGAAAGTTTCACTCTGCTGGGGAATACGGCAAGCTATCACT 1522  
2601 CGCTTCTCTCAGAAACCAAGAAAGTTTCACTCTCTGCGGGAAGCATGCAAGCTCTCGCT 2660  
1523 GCAGGACGTGTGGAGATGAAAGTAGAGGATTCGCACTGGCTCCGCGAGCGCCGGG 1582  
2661 GCAGGAGCTGAGCTGGAGATGAGCTGTGGGCTGCGCTTGGCTGCGAGAGGCGCAGG 2720  
1583 GAAGGACCGTGTCCCGCTGAGAGCACCTCTGAGGAGAGAGATCTCTGGTGAAGTTCTCT 1642  
2721 GGTGGCTGTGTCCGCGCGCAGAGCACCTCTGCTGAGGAGATCTCTGGCAGAGTTCTCT 2780  
1643 GTTCTGGCTGTGAGACATACGTGGTACAGCTGCTTAGTCACTTTTAAATCAGACAGA 1702

DG 2781 GCACTGGCTGATGAGTGTGTAGCTGCTGAGCTGCTCAGGTCTTTCTTTTATGTACACGA 2840  
QY 1703 GAGCATTTCAGAGAGACAGGCTCTTCTCTACCGTAAAGGTGTGTGTGAGCAAGCTCA 1762  
DB 2841 GACCATTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTCA 2900  
QY 1763 GAGCATTGAGTTCAGGCAACCTTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
DB 2901 AAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTGGAGACGA 2960  
QY 1823 GGTCAAGCATCAGCAGGACACTGCTAGCATGCGCCCATCTGAGACTGCGCTTCTATCCC 1882  
DB 2961 GGTCAAGCATCAGCAGGACACTGCTAGCATGCGCCCATCTGAGACTGCGCTTCTATCCC 3020  
QY 1883 CAAGCCCAAGGCTCTGGGCGCATTTGAAACATGAGTATAGCATGGGTACAGAGCTTT 1942  
DB 3021 CAAGCTGAGCGGCTGGGCGCATTTGAAACATGAGTATAGCTGCGGGAGCGAGAACGTT 3080  
QY 1943 GGCAGAGAAAGCAGGCGCCAGCATTTCAACAGCGCTCTCAAGACTCTCTTCAGCATGCT 2002  
DB 3081 CCGCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGGCACTGTTTCAGCGTGT 3140  
QY 2003 CAATATGAGCGGACAAACATCTCTACCTATGGGCTCTCTGACTGGGTATGAAGA 2062  
DB 3141 CAATGAGAGCGGCGCGCGCTCTCTGGGCTCTCTGGGCTCTGGAGCTGGAGCA 3200  
QY 2063 CATCTACAGGACTGGGCGCTTGTGTGGTGGTGGTGGTCTGGAGCAACACCCAG 2122  
DB 3201 TATCCACAGGCGCTGGGCGCATCTCTGCTGGTGGTGGGCGCGCCAGAGCCCGCCCTGA 3260  
QY 2123 GATGTACTTTTGAAGGAGAGTGTGACCGGGGCTATGATGTCATCCCCCAGGGTAAAGT 2182  
DB 3261 GCTGTACTTTTGAAGGAGAGTGTGACCGGGGCTATGATGTCATCCCCCAGGGTAAAGT 3320  
QY 2183 GGTGAGGTTTGTGCAATATGATCAGGCACTCGAGAGCACTGACTGTATCGGCGAGTA 2242  
DB 3321 CAGGAGGTTTATCGGCGAGCATCTCAAAACCC---AGAACAGCTCTGCTGGTGGTGGTA 3377  
QY 2243 TGCAGTGGTTCGAGAGATAGTCAAGGCGCAAGTCCAAAGTCTTTTAGGAGACAGGTCA 2302  
DB 3378 TGCGGTGTTCAGAGAGCGCGCCATCGGCGCACTCGGCAAGGCGCTTCAAGAGCAAGTCTC 3437  
QY 2303 CACCTCTCTGAGCTTCAGGCACTACATGCGGCGAGTCTTAAAGCATCTGAGAGATTCAGA 2362  
DB 3438 TACTTGAAGAGACTTCAAGCGTATGAGCAGTGTGCTGAGTCACTGAG---GA 3491  
QY 2363 TGCAGTGCACTGAGGAACTCGGTTGTCTGAGAGAGAGCATCTCTATGAATGAGAGAG 2422  
DB 3492 GACCAAGCGCTGAGGAGTGGCGTCTCATGAGAGAGTCTCTCTCTGAAATGAGGCGAG 3551  
QY 2423 CAGCAGCTGTGTGAGTCTTCTCTGCACTCTCTGCGCTCAAGTGTCTGAAAGATTTGTA 2482  
DB 3552 CAGTGGCTCTTTCAGAGCTCTTCTTACGCTTCTGCGCTTCAAGCGCGGTGCGCATCAGGG 3611  
QY 2483 CAGGTCTATAGCGAGTGGCGAGGCTCCCGCGAGGCTCCAGGCTATCCAGCTGCTCTG 2542  
DB 3612 CAGTCTTACGTTCCAGTGGCGAGGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 3671  
QY 2543 CAGTCTGTGTTCGAGACATGAGAACTGAGTGTGCTGAGGTGCGAGGAGTGGGT 2602  
DB 3672 CAGCTGTCTACGCGGACATGAGAACTGAGTGTGCGGGGATTCGGGCGGAGCGGCT 3731  
QY 2603 GCTTTTACGTTTGTGATGATCTTCTGTTGGTGGAGCTCTCACTTGGAGCAAGCAAAAC 2662  
DB 3732 GCTCTGCGTTGTGGTATGATTTCTTGTGTTGAGTCACTCTCACTCAGCGCGGAAAC 3791  
QY 2663 CTTCTCTACGACCTCTGCTTCCATGGCGTCTCTGAGTATGGGTGATGATTAACCTTCAGAA 2722  
DB 3792 CTTCTCTCAGGACCTCTGGTCCGAGGTGCTCTGAGTATGGCTGGTGGTGAATTCGGGA 3851  
QY 2723 GACAGTGTGAACTTCTCTCTGAGGCTCTGTAACCTGGGTGGTGGAGCTTCCATACAGCT 2782  
DB 3852 GACAGTGTGAACTTCTCTCTGAGGAGCGAGGCGCTGGGTGGCGAGCTTTTGTTCAGAT 3911









1643 GTTCTGGCTGATGGAACACATACGTGGTACAGCTGCTTAGGTGATCTTTTACATACACAGA 1702  
1690 GCACCTGGCTGATGAGTGTAGCTGCTGAGCTGCTGAGTCTTCTTTTATGTACCGGA 1749  
1703 GAGCACAATCCAGAGAAACAGGCTCTTCTTACCGTAAAGAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCAGCTTCAAAAGAAACAGGCTCTTCTTACCGGAAGAGTGTGTGGAGCAAGTGTGA 1809  
1763 GAGCATTGGAGTCAGCAACACCTTGAAGAGTGTGGTACCGAGAGTGTCAACAAGGGA 1822  
1810 AAGCATGGAAATCAGACACACTTGAAGAGGAGTGTGAGTGTGGAGTGTGGAAGCAGA 1869  
1823 GGTACAGCATCACAGAGACACTGTGGCTAGCATGCCCATCTGCAGAGTGTGGCTTATCCC 1882  
1870 GGTACAGCATCGGGAAGCCAGGCGCGCTGTGAGCTCCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAAGCGCTCGGCCCATTTGTGAACATGAGTTATAGCATGGTACAGAGCTTT 1942  
1930 CAAGCCCTGAGCGGCTCGGCCGATTTGTGAACATGAGTTATAGCATGGTACAGAGCTTT 1989  
1943 GGGCAGAAGGAAGCAGGCGCCAGCATTTCAAGAGTGTCTCAAGAGTGTCTTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCCAGGCTCTCACTCGAGGTGAAGCACTGTTTCAAGGCTGT 2049  
2003 CAATATGAGCGGACAAAACATCTTCACTTATGGGCTTCTGTACTGGTATGATGA 2062  
2050 CAATACGAGCGGCGCGCGCGCGCTCTGTGGCGGCTCTGTGTGGCGCTGTGACGA 2109  
2063 CATCTACAGAGCTGGCGGCGCTTTGTGTGGTGTGGTGTCTGTGGACAGACACCCAG 2122  
2110 TATCCACAGGCGCTGCGCACCTTGTGTGGTGTGGTGTGGCGCCAGGACCGCGCGCTGA 2169  
2123 GATGTACTTTTGAAGCAGATGTGACCGGCGCTATGATGCGATCCCGCAGAGTAAAGCT 2182  
2170 GCTGTACTTTTGAAGGTGATGTGAGCGGCGCTGTGACACCATCTCCCGCAGGACAGGCT 2229  
2183 GGTGGAGGTTTGTCCAAATATGATCAGGCACTCGGAGAGCAGTCTGTATCGGCCAGTA 2242  
2230 CACGAGGTATCGCCAGCATCATCA--AACCCAGAAACAGTACTGTGGTGGTGGTA 2286  
2243 TGAGTGTTCGAGAGATAGCAGAGCCAGTCCACAAAGTCTTTAGGAGACAGTCCAC 2302  
2287 TGCGGTGTTCAGAAAGCGCCCATCGGCGCTGTGGCGGCTTCAAGAGCCACAGTCTC 2346  
2303 CACCTCTGTGACTCCAGCCATACATGAGGCGAGTTCCTTAAAGCATCTGCAGGATTCAGA 2362  
2347 TACCTTGACAGACTCCAGCGGTACATGCGACAGTTCTGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGCATCGAGAACTCGGTTGTATCGAGCAGAGCATCTCTATGATGAGAGCAG 2422  
2401 GACAGGCCCTGAGGATGCGCTGTATCATCGAGCAGAGTCTCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCTCTTTGACTTCTTCTGACATCTCTGCGTCAAGTGTGCTGAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTCAGAGTCTTCTAGCTTCTATGTCACCAACGCGGTGCGCATCAGGG 2520  
2483 CAGTGTCTATACGAGTCCAGGCACTCCCGAGGCTCCAGGCTCCAGCTTCCACCTGCTCTG 2542  
2521 CAGTCTCTAGCTCAGTTCAGGAGATCCCGAGGCTCCAGCTTCTTCCAGCTGCTCTG 2580  
2543 CAGTCTCTTTTCGAGACATGAGAAACAGTGTGTCTGAGTGTGACGCGGATGGGTT 2602  
2581 CAGCCTGTGTACGGGACATGAGAAACAGTGTGTGCGGATTTGCGGCGGACGCGGT 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGTGTGAGCTCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTCGGTTTGTGATGATCTTCTGTGTGATGATCTCACTTCCACCTCACCGCAAAAC 2700  
2663 CTTCTCAGCACCTGTGTCAGCGCTTCTCAGTATGGTGTGATGATTAATCTTGCAAGA 2722  
2701 CTTCTCAGCACCTGTGTCAGCGGCTCTCAGTATGCTGCGTGTGATGCTGCGGATTCGCGAA 2760

QY 2723 GACAGTGTGAACCTCCCTGTGGAGCCTGGTACCTGGTGGTGCAGCTCCATACACAGCT 2782  
Db 2761 GACAGTGTGAACCTCCCTGTGAAGAGAGAGGCGCTGGTGGCAGCGCTTTTGTTCAGAT 2820  
QY 2783 GCCTGTCTACTGCTGCTTTTCCCTGGTGTGGTGTGTGCTGACACTCAGACTTTTGGAGGT 2842  
Db 2821 GCCGCCACCGGCTATTCCCTGGTGGGCTGTGTGTGATACCGGACCTCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACCTTCCAGAG 2902  
Db 2881 GCAGAGCAGTACTTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAAGCTGGGAAGACCATCGGAACAAGTCTCTGTGCTTGGGTTGAAGTG 2962  
Db 2941 CGGCTTCAAAGCTGGGAGGAACATCGCTCGAAACTCTTTTGGGCTTGGGCTGAAGTG 3000  
QY 2963 TCAGGCTTATTTCTAGACTTCAGGTGAACAGCTCCAGCTCCAGACAGTCTGCATCAATATA 3022  
Db 3001 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCTCCAGAGCGTGTGACCAACATCTA 3060  
QY 3023 CAAAGTCTTCTGCTTCAAGGCTACAGGTTCCATGCAATGTGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAAGTCTTCTGCTGAGCGGTACAGGTTTCAAGCATGTGTGCTGAGCTCCCATTTCA 3120  
QY 3083 CCAGCGTGTAGGAGAACCTCACATTTCTTGGGATCATCTCCAGCCAAAGCATCTTG 3142  
Db 3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTGGCGTCACTCTGACAGCGCTCCCT 3180  
QY 3143 CTGCTATGCTATCTGAAGGTCAAGAATCGAAGATGACATAAG-----GCCTC 3193  
Db 3181 CTGCTACTCCATCTGAAAGCCAAAGAACGAGGATGTGCTGGGCGCCAAAGGCGCGCC 3240  
QY 3194 TGCTCTTCTTCTGCTGAGCGCGCATTTGGCTGTCTGCTACAGGCTTCTGCTCAAGCT 3253  
Db 3241 CGGCTCTTCTGCTGAGCGCGCTGAGTGTGCTGTGCTGTGCTGTGCTGTGCTCAAGCT 3300  
QY 3254 GGCTGTCTATCTGTCTCATCTCAAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
Db 3301 GACTGCACACCGGTCTACCTAGCTGCCATCTCTGGGTCACCTCAGGACAGCCAGACGCA 3360  
QY 3314 GCTGTGCCGAGCTCCAGAGCGGACATGACCATCTTAAAGCTGACGCTGACCTGACCCAGC 3373  
Db 3361 GCTGAGTCCGAGCTCCCGGAGCAGCGCTGACTGCTGCTGAGGCGCGCAGCCACCCGCG 3420  
QY 3374 CCTAAGCAGACAGCTTTCAGACCATTTTGAATAA 3407  
Db 3421 ACTGCCCTCAGACTTTCAGACCATCTCTGACTGA 3454

## RESULT 8

AR182221 LOCUS AR182221 4015 bp DNA linear PAT 20-APR-2002  
DEFINITION Sequence 1 from patent US 637200.

ACCESSION AR182221

VERSION AR182221.1 GI:20225137

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 4015)

AUTHORS Morin, G.B.

TITLE Human telomerase catalytic subunit variants

JOURNAL Patent: US 637200-A 1 09-JAN-2002;

FEATURES Location/Qualifiers

source

1..4015

/organism="unknown"

BASE COUNT 663 a 1363 c 1275 g 714 t

ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGGCTTGAGCAAAATGACCCGGGCTCTCTGTTGCCCGCGGTGGCTCTCTGCTGG 82  
40 CCGGGCCACCCCGCGGATCCCGCGCTCCCGCTGCGAGCCGTGGCTCTCTGCTGG 99  
83 CAGCCGATACCGGGAGGTGGCGCTGGCAACCTTGTGGGGGCTGGGGCCCGAGGG 142  
100 CAGCCACTACCGGAGGTGCTCGCTGGCCACGCTTCTGGCGGCTGGGGCCCGAGG 159  
143 CAGGGGCTTGTGCAACCCGGGACCGGAAGATCTACCGGACTTGTGGTGGCCGATGCT 202  
160 CTGGCGGCTGGTGCAGCGGGGACCCCGCGCTTTCGCGCGCTGGTGGCCGATGCT 219  
203 AGTGTGATGCTACTGGGGCTCAGAGCTCCACTCGCGACCTTCTTCACACAGGTGTC 262  
220 GGTGTGGCTGGCTGGGACGCAAGCGCGCGCCCGCGCGCTCTCTTCGCGAGGTGTC 279  
263 ATCCCTGAAGAGCTGGTGGCCAGGGTGTGAGAGACTCTGCGAGGCGCAACGAGAGAA 322  
280 CTGGCTGAAGAGCTGGTGGCCGAGGTGCTGAGAGGCTGTGCGAGCGCGCGGAGAA 339  
323 CGTGTGGCTTTGGCTTTGAGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTT 382  
340 CGTGTGGCTTTGGCTTTGAGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTT 399  
383 CACTAGTAGCTGCTAGTACTTGCACCACTGTTATTGAGACCTCTGCTGCTGCTG 442  
400 CACCACAGCTGCGCAGCTACTGCGCCAAACAGGTGACCGAGCACTGCGGGGAGCGG 459  
443 TGCAATGATGCTACTTGTAGCGGAGTGGCGAGCACTGCTGCTGCTGCTGCTGCTG 502  
460 GCGGTGGGGGCTGCTGCTGCGCGGTGGCGAGCACTGCTGCTGCTGCTGCTGCTG 519  
503 CTGTGCTCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562  
520 CTGCGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAAAATTTGTGCCACCAAGGATATCTGGGCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCAAGCTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639  
623 CCGACCGCTGGGAGGAAATTTCACTAACTTAGTGTCTTACAAAGATCAAGAGCAGTAG 682  
640 GCTCTGGGATGC-----GAAAGGCTGGAACCATAG 672  
683 TCGCCAGGAGCAACCGAATCCCTGGCTTGGCTTCTGAGTACAAAGAGGATCTGAG 742  
673 CGTCAGGAGGCGGGGCTCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732  
743 TCTCACAGTACAGTGTGCTTACGTAAGAGGCGAGATGCTATCTGCTCCGAGAGT 802  
733 CAGTGCCAGCGGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
803 GGAGGAGGAGCCCAAGGAGGCTGCTACCAACCCCAATCAGGAGGAGGAGGAGGAGG 862  
793 GGAGGAGGAGCCCGTTGGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852  
863 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
914 AGGAAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972  
971 CTCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCCCCATCCACATCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1032  
1028 TGAGACAGACATTTCTTACTTCCAGGGGAGATGGCCAGAGGCTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTTACTTCTTACGGCA---CAAGGAGGAGCTGCGGCTCTCT 1089  
1088 CCTACTAGCAACCTCCAGCTAACTTGAATGCTGGGCGCAGGAGACTGGTGGAGATCATCT 1147

1090 CCTACTAGCTCTCTGAGGCCCGAGCTGACTGGGCTGGAGGCTGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAAGGCTTAGGACATCAGGACCACTCTGAGGACACACGCTTATCGGCTG 1207  
1150 TCTGGGTTCCAGGCTTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCCAGCG 1209  
1208 ATACTGGGAGATGGGCGCTTGTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
1210 CTACTGGCAATGGGCGCTTGTCTGAGGCTGCTTGGGAAACACGCGAGTGGCCCTA 1269  
1268 TGTCAAGACTCTCTCAGGCTCACTTGCAGGTTTCCAAACAGCAACCAACAGGTGACATGC 1327  
1270 CGGGGTGCTCTCAAGAGGCACTGCCGCTGCGAGCTGCGTCACTCCACAGCGCGGTGT 1329  
1328 CT-----TGAAACACAGCC 1342  
1330 CTGTGCCCGGAGAGCCCGAGGCTCTGTGGGCGCCCCCGAGGAGGAGGACACAGACC 1389  
1343 ACCGCACTCATGATTTGCTCCGCTGCACAGAGTCCCTGGCAGGTATATGCTTTCT 1402  
1390 CGTGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
1403 TCGGCTGCTCTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462  
1450 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509  
1463 CCGCTCTTTAAGAACTTAAAGAACTTCACTCTGCTGGGAAATACGGCAAGCTATCACT 1522  
1510 CCGCTCTCTCAGAAACACCAAGAACTTCACTCTGCTGGGAAAGATGCCAAGCTCTGCT 1569  
1523 GCAGAACTGATGTGAGAGATGAAAGTAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
1570 GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529  
1583 GAGAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642  
1630 GGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689  
1643 GTTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702  
1690 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749  
1703 GAGACATTCAGAAAGACAGCTCTTCTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1762  
1750 GACCACTGTTCAAAGACAGGCTCTTTCTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCT 1809  
1763 GAGCATTGGAGTCAAGCAACACCTTGAAGAGTCCGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGAAATCAGACAGCACTTGAAGAGGTCAGCTGCGGAGCTGTCCGAGCAGA 1869  
1823 GGTCAAGCATCACAGGACACCTGGTAGCCATGCCATCTGCAAGCTGCTGCTTCTATCC 1882  
1870 GGTCAAGCATCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1929  
1883 CAGGCGCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942  
1930 CAAGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989  
1943 GGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2002  
1990 CCGAGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2049  
2003 GAACTATAGCGGACAAACATCTCTCAGCTTATGGGCTCTTCTGCTGCTGCTGCTGCTGCT 2062  
2050 CAACTATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2109  
2063 CATCTACAGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2122  
2110 TATCTACAGGAGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2169  
2123 GATGATCTTTGTTAAGGAGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2182

2170 GCTGTACTTTGTCAAGTGGATGTGACGGGCGGTACGACACCAATCCCCAGGACAGCT 2229  
2183 GGTGAGGTTGTTGCCAATATGATCAGGACCTCGAGAGACAGTACTGTATPCGCCAGTA 2242  
2230 CACGGAGGTGATGCCAGCATCATCA---AACCCAGAAACAGTACTGCGTGCCTGGA 2286  
2243 TGCAAGTGTGCGGAGATAGCCCAAGCTCCACAAGTCTTTTAGAGACAGGTAC 2302  
2287 TGCGTGTGTCAGAGGCCCCATGCGGACGCTCGCAAGGCTTCAAGAGCCAGTCTC 2346  
2303 CACCTCTCTGACTCCAGCCATACATGAGGCGAGTCTTAAAGCATCTGCAAGTTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGTCACTGAGGAATCCGTTGTCTCATCGAGCAGAGCATCTCTATGAATGAGCAG 2422  
2401 GACGAGCCCGCTGAGGATGCGCTGTCTCATCGAGCAGAGCTCTCTCTCATGAGGCGAG 2460  
2423 CAGCAGCTGTTTGGACATCTCTGACATCTCTGCGTGCAGTGTCTGTAAGATTTGTTGA 2482  
2461 CAGTGGGCTCTTTCAGCTCTCTCTGAGTCTCTGAGTGTGCGACACGCGCTGCGCATCAGGG 2520  
2483 CAGTGTCTATACGAGTGCAGGCGCATCCCGAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCAGTGCAGGAGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTGGAGACATGAGAACAAAGCTGTTTGTGAGGTGCGAGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGCACATGAGAACAAAGTGTGTTGGGGGATTCGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTCTTGCATGACTTCTGTTGGTGAAGCTCTCATCTTGAGAACGAAAGAAC 2662  
2641 GCTCTGCTGTTGTTGATGATTTCTGTTGTTGATGACACCTCATCTCACCGAGGAAAC 2700  
2663 CTCTCTCAGACCTGCTGCATGCGGTTCTCTGATGATGAGTGCATGATAAATCTGCGAA 2722  
2701 CTCTCTCAGACCTGCTGCAGGTTCTCTGATGATGAGTGCATGATAAATCTGCGAA 2760  
2723 GACAGTGTGAACTTCTCTGAGGCTGCTACCTCGGTTGGTGCAGCTCCATCAGCT 2782  
2761 GACAGTGTGAACTTCTCTGAGGCTGCTACCTCGGTTGGTGCAGCTCCATCAGCT 2820  
2783 GCCTGCTCATGCTCTTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCACGCGCTATCTCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGCTACTCTCAGGTTATGCGGAGCTCAATTAAGAGAGGCTCACTTCCAGAG 2902  
2881 GCAGAGGACTACTCTCAGCTATGCGGAGCTCTCATCAGAGCAGTCTCATCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGAGACCATGCGGACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGGAGAGCAATGCGTGCAGAACTCTTGGGGTCTTGGGGTGAAGTG 3000  
2963 TCAGGCTTATTTCTAGACTTGCAGGTGAACAGCTTCCAGAGCTGCTGCATCAATATA 3022  
3001 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCTTCCAGAGCTGCTGCATCAATATA 3060  
3023 CAGATCTTCTGCTTCAAGCTTCAAGTTCATGATGATGATGATGATGATGATGATGATGATGAT 3082  
3061 CAAGATCTCTGCTGAGGCGTACAGGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAGGTTTCA 3120  
3083 CCAGGCTTATAGGAAGCTCTCATTTCTTCTGGGCTCATCTCCAGCCAGGATCTCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGCGGCTCTCT 3180  
3143 CTGCTATCTCTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 3193  
3181 CTGCTACTCTCTGAGAACCCAGAACCGAGGATGCTGCTGGGGGCAAGGGGCGCGC 3240  
3194 TGCT 3253  
3241 CGGCGCTCTGCGCTCCGAGGCGGTGCAAGTGGCTGTGCCAACAGGATCTCTCTCTCTCTCTCTCT 3300

QY 3254 GGCTGCTATTTCTGTCAATCTACAAATGTCCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
Db 3301 GACTCGACACCGTGTCACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
QY 3314 GCTGTGCGGAGAGCTCCAGAGCGCAATGACCAATCTTAAAGTGTGAGTGTGAGTGTGAGTGTG 3373  
Db 3361 GCTGAGTGTGAGAGCTCCCGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
QY 3374 CTTAAGCAGACAGCTTTCAGACCATTTTGTGACTAA 3407  
Db 3421 ACTGCGCTCAGCTTCAAGACCATCTGAGCTGA 3454

RESULT 9  
AR224455  
LOCUS AR224455 4015 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 1 from patent US 6440735.  
ACCESSION AR224455  
VERSION AR224455.1 GI:23333293  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Gaeta,F.C.A.  
TITLE Dendritic cell vaccine containing telomerase reverse transcriptase for the treatment of cancer  
JOURNAL Patent: US 6440735-A 1 27-AUG-2002;  
FEATURES Location/Qualifiers  
source 1..4015  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN

Query Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Best Local Similarity 58.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCCTGGCTTGAGCACATGACCGCGCTCTCTGTTGCCCGGGTGCCTCTCTCTGCTGCG 82  
Db 40 CCCTGGCACCCCGCGATGCGCGCTCTCTGCTGCGAGCGCTGCTCTCTCTGCTGCG 99  
QY 83 CAGCGATACCGGAGGTGTGCGCTGTGCAACCTTTGTGCGGCGCTTGGGGCCCGAGG 142  
Db 100 CAGCAGTACCGGAGGTGTGCGCTGTGCGCGCTGCTGCGGCGCTTGGGGCCCGAGG 159  
QY 143 CAGCGCTTGTGCAACCGCGGAGCCGAGATCTACCGCACTTGTGCTGCGCAATGCT 202  
Db 160 CTGCGGCTGTGTCAGCGCGGAGCCCGCGCTTTCGCGCTGTGCGCGAGTGCCT 219  
QY 203 AGTGTGATGACTGGGCTCAGAGCTTCACTGCGGACCTTCTCTTCCACAGGTGTC 262  
Db 220 GGTGTGCTGCTTGGGAGCGACGCGCGCTTCTCTGCGGCGCTTCTCTCTGCGAGGTGTC 279  
QY 263 ATCCCTGAAGAGCTGTGCGCAGGCTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTGTAGAGAGCTGTGCGCGCTGTGCTGAGAGCTGTGCGAGCGCGCGCGAGAA 339  
QY 323 GGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 382  
Db 340 GGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 399  
QY 383 CACTAGTAGGCTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 442  
Db 400 CACCCAGAGCTGTGCGAGCTTACCTGCCAACACGCTGACCGAGCTGCGGGGAGCGG 459  
QY 443 TGCATGATGCTACTGTTGAGCGAGTGTGCGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 502  
Db 460 GGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 519  
QY 503 CTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 562









1028 TGAGACACAGACATTTCTCTTACTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCATTT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGGCCCTCTTT 1089  
1088 CCTACTCAGCAACCTCCAGCCTAACTTTGACTGGGGCCAGGAGACTGGTGAGATCATCTT 1147  
1090 CCTACTCAGCTCTCTGAGGCCAGCTGACTGGGCTCGAGGCTCGTGAGACCACATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGAACCACTCTTGACGACACACCGTCTATCGGTCG 1207  
1150 TCTGGGTTCCAGGGCCCTGGATGCGCAGGCACTCCCGCAGGTTGGCCGCTGCGCCAGCG 1209  
1208 ATACTGGCAGATGGGCCCTGTTTCAACAGCTGCTGGTGAAACATGACAGAGTGCCCAATA 1267  
1210 CTACTGGCAATGCGGCCCTTGTTCTGAGAGTGTCTTGGGACACACGGCAGTGCCCCCTA 1269  
1268 TGTCAAGCTCTCAGGTCACANTCCAGGTTTCGAAACAGCAAAACCAAGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGAGCGCACTGCCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGACACACAGCCC 1342  
1330 CTGTGCGCGGAGAGGCCCTCTGTGGGCGCTCTGTGGGCGCCCGAGGAGGAGACACAGCCC 1389  
1343 ACCGCACCTCATGATTTGCTCCGCCCTGCAAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CGGTGCGCTGTGACGCTGCTCCGCCAGCAGCAGCAGCCCTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGGCTGTCTTGCAAGTGTGTCTGTAGTCTCTGGGTACAGGCAACATGAGCG 1462  
1450 GCGGGCTGCTGCGCGGCTGTGTGCCCCAGGCTCTGGGGTCCAGGCACAAACGAAAG 1509  
1463 CCGCTTTCTTAAGAACTTAAAGAGTTTCACTCTGTGGGAAATACGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAAAGTTTCACTCTCTGGGGAAGCAGTCTCGCT 1569  
1523 GCAGGAACTGATGTGGAAGATGAAGTAGAGGATGCGCACTGGCTCCGACAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGCAGGACTGGCGCTGGCTCGCAGGAGCCAGG 1629  
1583 GAAGGACGCTGCTCCCGCTGCAGAGCAGCGCTGTGAGGAGAGGATCTGCTCACTGCTCT 1642  
1630 GGTGGCTGTGTTTCGGCGCGAGAGCAGCGTCTGCTGAGGAGATCTGCGCAAGTCTCT 1689  
1643 GTTCTGGCTGATGACACATACGTGGTACAGCTGCTTATAGTCTATCTTTTACATCAGAG 1702  
1690 GCACTGGCTGATGATGTGTACGTGCTCGAGCTGCTCAGGCTCTTTCTTTTATGTACGGA 1749  
1703 GAGCATTCCAGAGACAGGCTCTTCTTACCGTAAAGTGTGTGGAGCAGCTGCA 1762  
1750 GACCAAGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTGTGAGCAAGTTGCA 1809  
1763 GAGCAATTGAGTCAAGCAACCTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCAATTGAAATCAGACAGCACTTGAAGAGGTGCAGCTCGGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCAGCAGACAGCTGGCTAGCCATGCGCCATCTGACAGCTGGCTTCTATCCC 1882  
1870 GGTCAAGCAGCATCGGGAAGCAGGCGCCGCTGCTGAGCTCCAGACTCGGCTTCACTCCC 1929  
1883 CAAGCCAAAGCCTGCGGCCCTATGTGAAATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAAGCTGACGGCTGCGGCCGATGTGTGAACATGGACTAGTCTGTGGGAGCCAGAGCTT 1989  
1943 GGGCAGAAAGGAGCGGCCAGCATTTCAACCGAGCTCTCAAGACTCTCTTCAAGCATGCT 2002  
1990 CCGCAGAGAAAAGAGGCGGAGCGTCTCACTCGAGGTTGAAGCACTGTTCAAGCTGCT 2049  
2003 CAAGCTATGAGCGGCAAAACATCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAAGTACAGCGGGCGCGGCCCTCTGGGCGCTCTGTGCTGGGCTGGAACA 2109

2063 CATCTACAGGACCTGGCGGCGCTTTGTGCTGCGTGTGCGTCTGTGACACACACCCAG 2122  
2110 TATCCACAGGGCTGCGCACCTTCTGTGCTGTGTGGGGCCAGGACCCGCGCCCTGA 2169  
2123 GATGTACTTTTAAAGCAGATGTGACCGGGCTTATGATGCAATCCGCCAGGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAAGTGGATGTGACGGGCGGTACGACACCATCCGCCAGGACAGCT 2229  
2183 GGTGGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CAGGAGGTCAATCGCAGCATCATCA---AAGCCAGAAACAGTACTGCGTGTGCGTGA 2286  
2243 TGCAAGTGGTCCGAGAGATAGCAAGGCCAAGTCCACAAGTCTCTTATGAGACAGGTCAAC 2302  
2287 TGCGTGGTCCAGAAAGCGCCCATGGGCACTCGGCAAGGCTTCAAGAGCAGCTCTC 2346  
2303 CAGCTCTGTGAACCTCAGCAGCATACATGGGCGAGTTCCTTAAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCAGCGCTTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGCATCAGGAACTCCGTTGTCATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACAGGCCCTGAGGAGTGCCTGTCATCGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGTACTTCTTCTGCACTTCTGCTCAGCTGTCTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTTTCGACGCTCTTCTTACGCTTCTATGTCACACCGCGTGGCATCAGGGG 2520  
2483 CAGGTGCTATACGAGTGCAGGGCATCCCGAGGGCTCCAGCTTATCCACCTGCTCTG 2542  
2521 CAAAGTCTACGTCAAGTGCAGGGATCCCGAGGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCCGAGACATGAGAAACAGCTGTTTGTGAGGTGCACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACCGCGCATGAGGAAACAGCTGTTTGGGGGATTCGGCGGAGCGGGT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGTGAGCGCTCACTTGGAGCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGTGATGATTTCTTGTGTGACACCTCACTCACCCACGCGAAAC 2700  
2663 CTTCCTCAGCACCTGCTGTCATGCGGCTTCTGAGTATGGGTGCATGATAAACTTCAGAA 2722  
2701 CTTCCTCAGGACCTGCTGGTCCGAGGTGTCTCTGATATGGTCTGCTGGTGAATTCGGGA 2760  
2723 GACAGTGGTGAATCTCCCTGTGAGGCTTGGTACCTGTGGTGTGAGTCTCCATACAGCT 2782  
2761 GACAGTGGTGAATCTCCCTGTAGAGCAGGCGCTTGGTGGCACGGCTTTTGTTCAGAT 2820  
2783 GCTGTCTACCTGCTGTTTCCCTGTGCTGGTGTGCTGTGGACACTCAGACTTTTGGAGGT 2842  
2821 GCGGCCCGCGCTTATTTCCCTGTGCGGCTGTGCTGTGGATACCCGACCCCTGGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAG 2902  
2881 GCAGAGGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAAGACCATGGGAAACAGTCTCTGTGCTGCTTGGGTTGAAGTG 2962  
2941 CGGCTTCAAAGGCTGGGAGGAACATGGTTCGCAAACTCTTGGGGTCTTGGGCTGAAGTG 3000  
2963 TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCACAGCTGTTTCTGATTTGCAAGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA 3060  
3023 CAAGATCTCTCTGCTTACGGCTTACAGTTTCATGATGTGTGATTTCAGCTTCCCTTGA 3082  
3061 CAAGATCTCTCTGCTCAGCGCTTACAGGTTTTCACGATGTGTGCTGCACTCCCAATTTCA 3120  
3083 CCAGCGTGTGTAGGAAGAACTCACATTTCTTCTGGGCACTCATCTCCAGCAAGCATCTG 3142  
3121 TCAGCAAGTGTGGAAGAACCCACATTTTCTGCGGTCTCTCTGACAGGCTCTCCT 3180  
3143 CTGCTATGCTATCTCTGAAGGTCAAGAAATCCAGGAATGACACTAAAG-----GCCTC 3193

3181	CTGCTACTCCATCTCTAAAGCCAAAGAACCGAGGGATGTCGCTGGGGGCCAAGGGCGCCG	3240
3194	TGGCTTCCTTTTCTCTGAAGCCGACATTTGGCTCTGCTACCAAGGCTTTCTTGCTCAAGCT	3253
3241	CGGCCCTCTGCTCCCTCCAGAGCCGCTGCAGTGGCTGTGCCAACCAAGCATTTCTGCTCAAGCT	3300
3254	GGCTGCTCATTTCTGTCTATCTACAAATGTTCTCTGGGACCTCTGAGACAGCCGCAAAACT	3313
3301	GACTCGACACCGTGTCACTTACCTACGTGGCACCTCTCTGGGGTCACTCAGGACAGCCGACGCA	3360
3314	GCTGTGCCGGAAGCTCCACAGAGCGCAATAGACCATCTTAAAGCTGCAGCTGACCCAGC	3373
3361	GCTGAGTCGGAAGCTCCCGGGGACGACGTGCTGCCTGGAGGCGCAGCAACCCGCG	3420
3374	CCTAAGCAGACACTTCAGACCATTTTGGACTAA	3407
3421	ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA	3454
ULT 11		
43328		
US	AR243328	4015 bp mRNA linear PAT 20-DEC-2002
INITION	Sequence 1 from patent US 6475789.	
SSION	AR243328	
SION	AR243328.1	GI:27290539
WORDS		
ORGANISM	Unknown.	
	Unknown.	
ERENCE	1 (bases 1 to 4015)	
AUTHORS	Cech,T.R., Lingner,J., Nakamura,T., Chapman,K.B., Morin,G.B., Harley,C.B. and Andrews,W.H.	
TITLE	Human telomerase catalytic subunit: diagnostic and therapeutic methods	
JOURNAL	US 6475789-A 1 05-NOV-2002;	
ATURES	Location/Qualifiers	
source	1..4015	
	/organism="unknown"	
SEQUENCE	663 a 1363 c 1275 g 714 t	

QY	443	TGCATGATGCTACTGTGAGCGAGTGGGCGAGCAGCCTGTGTGTACTCTGTGGCACA	502
DB	460	GGCGTGGGGGCTGTCTGCGCGCGTGGGCGACGACGTGTGTGTTTACCCTGTGGCAG	519
QY	503	CTGTGCTCTTTATCTTCTGGTGCCGCCAGCTGTGCCCTACACAGGTGTGTGGTCTCCCT	562
DB	520	CTGCGGCTCTTTGTGCTGGTGGCTCCAGCTGCGCTACCAAGTGTGGGCGCGCT	579
QY	563	GTACAAATTTGTGCACACAGGATATCTGGGCTCTGTGTCCGCTAGTTACAGGCCAC	622
DB	580	GTACAGCTGGGCGCTGCCACTCAGCGCGGCGCCGCCACACGCTAGTGGACCCGGAAG	639
QY	623	CCGACCCGTGGGAGGAAATTTACATAACCTTAGTTCCTTCAACAGATCAAGAGCAGTAG	682
DB	640	CGTCTGGATGC-----GACGGGCTTGGAAACCATAG	672
QY	683	TCGCGAGGAGCACCGAAACCTTGGCTTGCATCTCGAGGTACAAGAGGCATCTGAG	742
DB	673	CGTCAAGGAGCGCGGGTCCCCCTGGGCTGCGACCCCGGGTGCAGAGAGCGCGGGG	732
QY	743	TCTCACAGTACAAGTGTGCCCTTCAGCTAAGAAGGCCAGATGCTATCTCTGCCAGAGT	802
DB	733	CAGTGCCAGCGAGTCTCGGCTGCCCAAGAGGCCAGGCGTGGCGCTGCCCTTGAGCC	792
QY	803	GGAGGAGGACCCCA CAGCGAGGTCTACCAACCCATCAGGCAATCATGGGTGCCAAG	862
DB	793	GGAGCGGACGCGGTTGGGAGGGGTCCTGGGCCCAACCGGCGAGGACGCTGGACCGAG	852
QY	863	TCCTGCTCGGTCCCGGAGTGCTACT-----GCAGAGAAAGATTTGTCTCTAA	913
DB	853	TGACCGTGGTTCTGTGTGTGTCACTGCCAGAGCCCGCAAGAACCACTCTTTGGA	912
QY	914	AGGAAAGGTGTGCACTGAGTCTTC---TGGGTGGTGTGTGTATAACACAGCCAG	970
DB	913	GGGTGGCTCTCTGGCAGCGGCACCTCCACCACCTCCGTGGGCCGCCAGACACAGCGGG	972
QY	971	CTCCACAT---CTGTGTGTCAACCCCGCCAAATGCTTTTTCAGTCTAGGCCATTAT	1027
DB	973	CCCCCATCCACATCGGGGCCACAGTCCCTGGGACACGCTTGTCCCGGGTGTACGC	1032
QY	1028	TGAGACGAGATTTCTTTTACTTCCAGGGGAGATGGCCAGAGCGTCTAAACCCCTCAT	1087
DB	1033	CGAGACGAGCATTTCTCTACTCTCTAGGCGA---CAAGAGACAGTGGGGCCCTCTT	1089
QY	1088	CTACTACGAACTCCAGCCTAACTTGACTGGGCCCAGAGACTGGTGGAGATCATCTT	1147
DB	1090	CCTACTAGCTCTCTGAGGCCAGCCTGACTGGCGCTGGAGGCTCGTGGAGACCATCTT	1149
QY	1148	TCTGGCTCAAGGCTAGGACATCAGGACCATCTCTGAGGACACACCGTCTATCGGTCG	1207
DB	1150	TCTGGTTCCAGGCCCTGGATGCGAGGACTCCCGCAGGTGGCCCGCTGCCCGACG	1209
QY	1208	ATACTGGCAGATCGGCCCTGTTCAAACAGTGTGTGTGGAACATGACAGAGTGGCAATA	1267
DB	1210	CTACTGGCAATGCGGCCCTTTCTGGAGTGTCTGGGAACACACGCGAGTGGCCCTA	1269
QY	1268	TGTCAGACTCTCTCAGGTCAATTCAGGTTTCGAAACAGCAACAAAGGTGACAGATGC	1327
DB	1270	CGGGGTGTCTTCAAGACGCACTGCCGCTGGAGCTGGGTCAACCCAGCAGCGGTGT	1329
QY	1328	CT-----TGAAACACAGCCC	1342
DB	1330	CTGTGCGCGGAGAGCCCGAGGGCTCTGTGGCGCCCCCGAGGAGGAGGACACAGACC	1389
QY	1343	ACCGCACCTCATGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTTCT	1402
DB	1390	CGGTGCTGCTGAGCTGTCCGCCAGCACAGCAGCCCTGGCAGGTGTACGCTTCGT	1449
QY	1403	TCGGGCTGTCTCTGCAAGGTGTGTCTAGTCTCTGGGGTACACGAGCAATGAGCG	1462
DB	1450	GGGGGCTTGCCTGGCGGGTGTGTGGCCCGCCAGGCTCTGGGGCTCCAGGCAACGAACG	1509



[REDACTED]

2003	CAACTATGAGCGGACAAAA	CATCCTCACCTTATGGGGTCTTCTGTACTGGGTATGATGA	2062
2050	CAACTGAGCGGGCGGGCGCCCGGCTCTGTGCTGGGCTGGACGA	2109	
2063	CATCTACAGGACCTGGGGGCGCTTGTGTGCTGGTGTGGTGTCTCTGGACACGACACCCAG	2122	
2110	TATCCACAGGCGCTGGGCGACCTTTCGTGTGCTGTGGGGCCAGGACCGCGCGCTCGA	2169	
2123	GATGTACTTTTAAAGCAGATGTGACCGGGGCTTATGATGCAATCCCCACAGGTTAAGCT	2182	
2170	GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCCACGACAGGCT	2229	
2193	GGTGGAGGTTGTTCGCAATGATCAGGCACCTCGAGAGCAGCTACTGTATTCGCCAGTTA	2242	
2230	CACGAGGTCAATGCCAGCATCATCA--AACCCAGAACAGTACTGCGGTGCGTGGTA	2286	
2243	TGCCAGTGTCTCGGAGAGATAGCCAAGGCCAAGTCCACAAGTCCCTTTAGGAGACAGGTCAC	2302	
2287	TGCCGTGTCCAGAAGCGGCCCATGGGCACGTCCGCAAGGCTTCAAGAGSCCAGCTCTC	2346	
2303	CACCTCTCTGACCTCAGGCCATACATGGGCCAGTTCCTTAAGCATCTGAGAGATTCCAGA	2362	
2347	TACCTTGACAGACCTCAGCGCTACATGCGCAGATTGCTGGCTCACCTGCAG-----GA	2400	
2363	TGCCAGTGCATCTGAGGAATCTCCGTGTCTCATCGAGCAGAGCATCTCATGAATGAGAGCAG	2422	
2401	GACCAGCCGCTGAGGGATGCCGTCTCATCTCGAGCAGAGCTCTCCCTGATGAGGCCAG	2460	
2423	CAGCAGCCTGTTGACATTCTTCTGACATCTCTGCGTACAGTGTGCTGAAGATTGTTGA	2482	
2461	CAGTGGCCCTCTTCACAGCTCTTCTCATGCTTCATGTGCCACACGCGGTGCGCATCAGGG	2520	
2483	CAGGTGCTATTACGACGTGCCAGGGCATCCCCAGGGCTCCAGGCTATCCACCCCTGCTCTG	2542	
2521	CAAGTCTTACGTCCAGTGCAGGGGATCCGCGAGGGCTCCATCTCTCCAGCTGCTCTG	2580	
2543	CAGTCTGTGTTTCGAGACATGAGAGAACAGCTGTTTCTGAGGTGTGACGCGGATGGGTT	2602	
2581	CAGCCTGTGCTACGCGCACATGAGAGAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT	2640	
2603	GCTTTTACGTTTTCTGATGACTTCTGTGTGTGAGCGCTCACITTTGGAACCAAGCAAAAAC	2662	
2641	GCTCTCGTTTGTGTGATGATTTCTTGTGTGACATCTCACCTCACCCACGCGAAAC	2700	
2663	CTTCTCAGACACCTGTGTCCATGGCGTTCTGTAGTATGGGTGCATGATAAATTGCGAGAA	2722	
2701	CTTCTCAGGACCCCTGTGTCCAGGTGTCCCTGAGTAGTGTGCTGCGTGTGAACTTTCGCGAA	2760	
2723	GACGTGTGAACCTTCCTGTGGAGCCTGTGTACCTGTGGTGTGTGAGCTCCATACCACTG	2782	
2761	GACAGTGTGAACCTTCCTGTAGAGACGAGGCCCTGGGTGSCACGGCTTTTGTTCAGAT	2820	
2783	GCTCTGCTCACTGCGCTGTTTCCCTGGTGTGCTGTGCTGTGACATCTCAGACTTTGGAGGT	2842	
2821	GCGGCCCAACGCGCCTATTCCCTGCTGCGGCGCTGCTGCTGTGATAGCGAGCTTTGTTGAGGT	2880	
2843	GTTCGTGCACTACTCAGTTATGCCAGACCTCAATTAAGACAGGCTCACCTTCCAGAG	2902	
2881	GCAGAGCGACTTCTCAGCTATGCCCGACCTTCATCAGAGCCAGTCTCACTTCAACCG	2940	
2903	TGCTTTCAAAGCTCGGAGACCATGTGGAAACAAGCTCTCTGTCGGTTCGCGTTGAAAGTG	2962	
2941	CGGCTTCAAGGCTGGGAGGAACATGCGCTCGCAAACTCTTTTGGGGTCTTTCGCGCTGAAAGTG	3000	
2963	TCAAGCTTATTTCTAGACTTTCAGAGGTGAACGCGCTCCAGACAGTCTGCATCAATATATA	3022	
3001	TCAAGCCTGTTTCTGGATTTTCAGGTGAACAGCCTTCCAGAGCGGTGTGCAACAATCTA	3060	
3023	CAAGATCTTCTGCTCAGGCCTACAGGTTCCATGATGTGTGATTCAGCTTCCCTTTGA	3082	
3061	CAAGATCTCTGTGTGAGGCTACAGGTTTTCAGCATGTGTGTGCTGCACTTCCATTTCA	3120	

QY	3083	CGACGGTGTAGGAAGAACTCATCTCTTTCTGGGCAATCATCTCCAGCCAGCAATCCTG	3144
Db <td>3121</td> <td>TACGCAAGTTTGGGAAGAACCCACATTTTCTGGCGGTCACTCTGACACGGGCTCCCT</td> <td>3180</td>	3121	TACGCAAGTTTGGGAAGAACCCACATTTTCTGGCGGTCACTCTGACACGGGCTCCCT	3180
QY <td>3143</td> <td>CTGCTATGCTATCTGAAGTCAAGAAATCCAGGAATGACACTAAAG-----GCCTC</td> <td>3193</td>	3143	CTGCTATGCTATCTGAAGTCAAGAAATCCAGGAATGACACTAAAG-----GCCTC	3193
Db <td>3181</td> <td>CTGCTATCTCAATCTTGAAGCCAGAAACGACGGATGTGCTGGGGCCAGGGCGCGC</td> <td>3240</td>	3181	CTGCTATCTCAATCTTGAAGCCAGAAACGACGGATGTGCTGGGGCCAGGGCGCGC	3240
QY <td>3194</td> <td>TGGCTCTCTTCTCTGAAGCCGCAATTTGGCTCTGTCTACAGAGCTTCTGTCTAAGCT</td> <td>3253</td>	3194	TGGCTCTCTTCTCTGAAGCCGCAATTTGGCTCTGTCTACAGAGCTTCTGTCTAAGCT	3253
Db <td>3241</td> <td>CGGCGCTCTGCGCTCCGAGGCCGTGAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT</td> <td>3300</td>	3241	CGGCGCTCTGCGCTCCGAGGCCGTGAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT	3300
QY <td>3254</td> <td>GSCTCTCATTTCTGTCTATCTACAAATGTCTCTGGGACCTCTGAGACAGCCCAAAACT</td> <td>3313</td>	3254	GSCTCTCATTTCTGTCTATCTACAAATGTCTCTGGGACCTCTGAGACAGCCCAAAACT	3313
Db <td>3301</td> <td>GACTCGACACCGTGTCACTTACGTGCCATCTCTGGGTCACTCAGACAGCCAGACGCA</td> <td>3360</td>	3301	GACTCGACACCGTGTCACTTACGTGCCATCTCTGGGTCACTCAGACAGCCAGACGCA	3360
QY <td>3314</td> <td>GCTGTGCGGAAGCTCCGAGCGCAATGACCACTCTTTAAAGCTGAGTGAACCGAGC</td> <td>3373</td>	3314	GCTGTGCGGAAGCTCCGAGCGCAATGACCACTCTTTAAAGCTGAGTGAACCGAGC	3373
Db <td>3361</td> <td>GCTAGTCCGAAGCTCCGCGGACGACGCTGACTGCTCTGGAGCGCGCAGCAACCCGCG</td> <td>3420</td>	3361	GCTAGTCCGAAGCTCCGCGGACGACGCTGACTGCTCTGGAGCGCGCAGCAACCCGCG	3420
QY <td>3374</td> <td>CCTAAGCACACACTTTTCAGACCAATTTTGACTAA</td> <td>3407</td>	3374	CCTAAGCACACACTTTTCAGACCAATTTTGACTAA	3407
Db <td>3421</td> <td>ACTGCCCTCAGACTTCAAGACCATCTGACTGA</td> <td>3454</td>	3421	ACTGCCCTCAGACTTCAAGACCATCTGACTGA	3454

  

RESULT 13	AR265996	Sequence 3 from patent US 6492171.	DNA	linear	PAT 10-APR-2003
LOCUS	AR265996		4015 bp		
DEFINITION	Sequence 3 from patent US 6492171.				
ACCESSION	AR265996				
VERSION	AR265996.1	GI:29694842			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				

  

REFERENCE 1 (bases 1 to 4015)	AUTHORS	Monla, B. P., Gaarde, W. A., Freier, S. M. and Wanciewicz, E.
TITLE	Antisense modulation of HERT expression	
JOURNAL	Patent: US 6492171-A 3 10-DEC-2002;	
FEATURES	Location/Qualifiers	
	1..4015	
	/organism="unknown"	
BASE COUNT	663 a 1363 c 1275 g	714 t
ORIGIN		

  

Query Match	45.8%;	Score 1600.8;	DB 6;	Length 4015;
Best Local Similarity	68.9%;	Pred. No. 0;		
Matches 2379;	Conservative 0;	Mismatches 967;	Indels 108;	Gaps 9;

  

QY	23	CCGCGCTTGAGCAATGACCCGCGCTCTCTGTTCCCGCGGTGCGCTCTCTGCTGCG	82
Db <td>40<th>CCGCGCACCCCGCGATCGCGGCTCCCGCTCGCGAGCGGTGCTCCCTGCTGCG</th><th>99</th></td>	40 <th>CCGCGCACCCCGCGATCGCGGCTCCCGCTCGCGAGCGGTGCTCCCTGCTGCG</th> <th>99</th>	CCGCGCACCCCGCGATCGCGGCTCCCGCTCGCGAGCGGTGCTCCCTGCTGCG	99
QY <th>83</th> <th>CAGCGATACCGGAGGTGTGGCGCTGTGGCAACCTTTGTGGCGGCTTGGGCGCGAGG</th> <th>142</th>	83	CAGCGATACCGGAGGTGTGGCGCTGTGGCAACCTTTGTGGCGGCTTGGGCGCGAGG	142
Db <th>100</th> <th>CAGCCACTACCGGAGGTGTGGCGCTGTGGCAACCTTTGTGGCGGCTTGGGCGCGAGG</th> <th>159</th>	100	CAGCCACTACCGGAGGTGTGGCGCTGTGGCAACCTTTGTGGCGGCTTGGGCGCGAGG	159
QY <th>143</th> <th>CAGCGGCTTGTGCAACCCCGGAGCCGGAAGATCTACCGCACTTTTGGTTGCCAATGCT</th> <th>202</th>	143	CAGCGGCTTGTGCAACCCCGGAGCCGGAAGATCTACCGCACTTTTGGTTGCCAATGCT	202
Db <th>160</th> <th>CTGGCGGCTGTGTGAGCGCGGGGACCGCGGCTTTCGCGGCTGTGGTGGCCAGTGCT</th> <th>219</th>	160	CTGGCGGCTGTGTGAGCGCGGGGACCGCGGCTTTCGCGGCTGTGGTGGCCAGTGCT	219
QY <th>203</th> <th>AGTGTGCAATGACTGGGGCTCAAGCTCTCACTGCGGACCTTTCTTCCACAGGTGTG</th> <th>262</th>	203	AGTGTGCAATGACTGGGGCTCAAGCTCTCACTGCGGACCTTTCTTCCACAGGTGTG	262
Db <th>220</th> <th>GGTGTGCGTCTGGACACGACGGCGCGCCCGCGCTCTCTTCCGCGAGGTGTC</th> <th>279</th>	220	GGTGTGCGTCTGGACACGACGGCGCGCCCGCGCTCTCTTCCGCGAGGTGTC	279
QY <th>263</th> <th>ATTCCTGAAGAGCTGTGTGGCCAGGGTGTGACAGACTCTCGAGCGCAACGAGAGAA</th> <th>322</th>	263	ATTCCTGAAGAGCTGTGTGGCCAGGGTGTGACAGACTCTCGAGCGCAACGAGAGAA	322
Db <th>280</th> <th>CTGCTCTGAAGAGCTGTGTGGCCAGGTGTGTGACAGGCTGTGAGCGCGCGCGAGAA</th> <th>339</th>	280	CTGCTCTGAAGAGCTGTGTGGCCAGGTGTGTGACAGGCTGTGAGCGCGCGCGAGAA	339
QY <th>323</th> <th>CGTGTGGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGAGCGGGCTTCCATGGCCTT</th> <th>382</th>	323	CGTGTGGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGAGCGGGCTTCCATGGCCTT	382
Db <th>340</th> <th>CGTGTGGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGAGCGGGCTTCCATGGCCTT</th> <th>399</th>	340	CGTGTGGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGAGCGGGCTTCCATGGCCTT	399

383 CACTAGTAGCGTGGTGTAGCTACTTGGCCAACTGTTATTAGAACCTCGTGTCAAGTGG 442  
400 CACCACAGCGTGGCAGCTACTTGGCCAACTGTTATTAGAACCTCGTGTCAAGTGG 459  
443 TGCATGATGCTACTGTTAGCGGAGTGGGCGGAGCTGCTGCTTACTGCTGCTGCA 502  
460 GCGTGGGGGCTGCTGCTGCGCGCGCTGCGGCGAGCTGCTGCTGCTGCTGCTGCA 519  
503 CTGTGCTCTTTATCTTCTGCTGCGCGCGCTGCGGCGAGCTGCTGCTGCTGCTGCT 562  
520 CTGCGCGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
563 GTACCAATTTGTCGACACGAGATCTGCGCGCTCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCAAGCTGCGCGCTGCTGCTGCGCGCGCTGCGGCGAGCTGCTGCTGCTGCTGCT 639  
623 CCGACCGCTGCGCGAGGATTTCACTACTTACTGCTTCTTCAACAGATCAAGAGCAGTAG 682  
640 CGCTGCGGATGC-----GACCGGCGCTGCGGAGCAGTAG 672  
683 TCGCCAGGAGCAACGAAACCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742  
673 CGTCAGGAGGCGCGGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732  
743 TCTCACAGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 802  
733 CAGTCCAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
803 GGAGGAGGAGCCCAACAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862  
793 GGAGCGGAGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852  
863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
914 AGGAAAGTGTCTGACCTGAGTCTCTC---TGGGTGCGTGTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGGTGGCTGCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972  
971 CTCACAT---CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCCCCCATCCACATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032  
1028 TGAGACCAAGATTTCTTACTCCAGGAGAGTGGCAAGAGCGCTGCTGCTGCTGCTGCTGCT 1087  
1033 CGAGACCAAGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1089  
1088 CCTACTGACCACTCTGAGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1147  
1090 CCTACTGAGCTCTCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
1148 TCTGCGCTCAAGGCTGAGCATCAGGACACTCTGAGGAGCACTGCTGCTGCTGCTGCTGCTGCT 1207  
1150 TCTGSGTTCCAGGCGCTGAGTGCAGAGGCTTCCCGAGGTTGCGCGCTGCTGCTGCTGCTGCT 1209  
1208 ATACTGGCAGATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
1210 CTACTGGCAATGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1269  
1268 TGTACAGTCTCTCAGTGCATTCGAGTTTTCGAGGAGTTCGAGGAGTTCGAGGAGTTCGAG 1327  
1270 CCGGTGTCTCTCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1329  
1328 CT-----TGAAACACAGCGCC 1342  
1330 CTGTGCGCGGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1389  
1343 ACCGACCTCATGAGTTTCTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402  
1390 CCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
1403 TCGGCGCTGCTCTGCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462

1450 GCGGCGCTGCTGCTGCGCGCTGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCG 1509  
1463 CCCTCTTTTAAAGACTTAAAGAGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1522  
1510 CCCTTCTTCAAGAGCAACCAAGAGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569  
1523 GCAGGAATGATGTGGAAGATGAAAGTAGAGATTTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
1570 GCAGGAGCTGAGCTGGAAGATGAGGCTGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1629  
1583 GAAGGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642  
1630 GGTGCT 1689  
1643 GTTCTGCT 1702  
1690 GCACTGCT 1749  
1703 GAGCACAATTCAGAGAAACAGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1762  
1750 GACCACTGCTTCAAGAGAAACAGGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1809  
1763 GAGCATTGAGTGCAGCAACACCTTGCAGAGAGTGGGCTTACCGGAGCTGCTCACAAGAGGA 1822  
1810 AAGCATTGAGTGCAGAGCACTTGCAGAGAGTGGGCTGCTGCGGAGCTGCTCGAGAGCAGA 1869  
1823 GGTGAGCATCACAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882  
1870 GGTGAGCAGCATCGGAGAAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929  
1883 CAAGCCCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942  
1930 CAAGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989  
1943 GGGCAGAGAGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2049  
2003 CAATATGAGCGAGCAAAACATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2062  
2050 CAATGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2109  
2063 CATCTAGAGACTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATGCTGCTTGTAAAGCGAGTGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2182  
2170 GCTGCTGCTTGTCAAGGTGAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2229  
2183 GGTGAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2242  
2230 CACGAGGCTCATCGCCAGCATCATCA---AACCAGAGCACTGCTGCTGCTGCTGCTGCTGCTGCT 2286  
2243 TGCATGCTGCGAGAGATAGCCAGGCGAGTTCACAGTCTTCTTAGGAGACAGGCTGCTGCTGCT 2302  
2287 TGCCTGCTTGTCAAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346  
2303 CACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2362  
2347 TACCTTGACAGACTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400  
2363 TGCAGTGCATGAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2422  
2401 GACGAGCGCTGAGGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
2423 CAGCAGCT 2482  
2461 CAGTGGCTGCTTGCAGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGTGTGCTATACGAGTGGCAGGAGTCCCGCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2542





793 GGAGCGAAGCGCGTGGCGAGGGTCTCTGGGCCCAACCGCGGCAGAGCGGTGACCGAG 852  
863 TCCTGCTCGGTCCCGAGGTCGCTACT-----GCAGAGAAAGATTGTCTCTTAA 913  
853 TGACCGTGGTTCTGTGTGTGTACCTGCGAGACCGCGGAGAGCCACTTTTGA 912  
914 AGAAAGGTGTCTGACCTGAGTCTCTC-----TGGGTGGTGTGTCTGTAACACAGCCGAG 970  
913 GGTGTGGCTCTCTGGACACGCGCACTCCCAACCATCCGTGGGCCCGCCAGCACACGCGGG 972  
971 CTCACAT---CTCTGCTGTCAACCAACCGCGCAAAATGCTTTACGCTTCAGCTCAGCCCATTTAT 1027  
973 CCCCCCATCAATCCGCGCACCAAGTCCCTTGGGACACGCTTGTCCCGGTGTAGGC 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAAGAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGGA---CAAGAGCAGCTGCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCCAGCTTAACTTTGACTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTCAGGCCCGAGCTGACTGGCGTCGGAGGCTCGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGACATCAGAACCACTCTGCAGGACACACCGCTCTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCG 1209  
1208 ATACTGGAGATCGGCGCTGTTCACAGCTGTGTGTGAACCATGACAGAGTGCATA 1267  
1210 CTACTGGCAATCGGCGCTGTCTTCTGGAGCTGTCTGGGAAACCAACGCGCAGTGCCTTA 1269  
1268 TGTCAAGACTCTCAGGTCAATTCAGAGTTTCGAAACAGCAAAACCAAGGTGAAGATGC 1327  
1270 CGGGGTCTCTCAAGACGACTGCCCGCTGGAGTGGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAACACAGGCC 1342  
1330 CTGTGCGCGGAGAACCCCGAGGCTCTGTGCGGCCCGCCGAGGAGGAGACACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CGTGGCTGGTGCAGTCTCCGACAGCAGACAGCCCTCGGAGGTGTACGGCTCGT 1449  
1403 TCGGGCTGTCTCTCAAGGTGTGTCTAGTCTCTGGGTACCAAGCAATGAGCG 1462  
1450 GCGGGCTGCTCGCGCGGTGTGTGCGCCAGGCTCTGGGCTCCAGGCAACAGAACG 1509  
1463 CCGCTCTTTAAGAACTTAAAGAGTTCTCTGTTGGGAAATACGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAGTTCTATCTCCCTGGGGAAGCATGCCAGCTCTGCT 1569  
1523 GCAGGAACATGATGTGAAGATGAAGTAGAGGATTGCCACTGGCTCCGACAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGACCTGGCTTGGCTGCCAGAGGCCGAGG 1629  
1583 GAAGGACGTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCTCTGGCTACGTTCT 1642  
1630 GGTGTGGTGTGTTCGGCGCGCAGAGCACCGTCTGGGTGAGGAGATCTCTGGCAAGTTCT 1689  
1643 GTTCTGGCTCATGACACATACCTGTGTACAGTGTCTTAGGTCAATCTTTTATACATCAGA 1702  
1690 GCACCTGGCTGATGAGTGTGTACGTCTGAGGTGTCTCAGTCTTTCTTTTATGACGGA 1749  
1703 GAGCAATTCAGAGAAAGAGGTCTTCTTACCGTAAAGGTGTGTGGAGAGAGCTGCA 1762  
1750 GACCAGTTTCAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGATTGGAGTCAGGCAACCTTCAGAGAGTGGCTACCGGAGCTGTACAGAGGA 1822  
1810 AAGATTGGATCAGACAGCACTTGAAGGGTGTGAGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCAACAGGACACTGGGTAGCCATGCCCATCTCGAGACTGGCTTCATCCC 1882  
1870 GGTCAAGGCTGGGAGGAACATGCGTGGCGCAAACTCTTTGGGGTCTTGGGCTGAAGTG 1929

1883 CAAGCCCAACGGCTCGCGGCCCAATTGTGAACATGAGTTATAGCATGGTACCAGAGCTTT 1942  
1930 CAAGCCTGACGGCTGCGGCGGATTGTGAACATGAGTACGTTGAGGAGCCAGAACGTT 1989  
1943 GGGCAGAGAGAGCAGCGCCAGCATTTTCAACCCAGCGTCTCAAGACTCTCTTTCAGCATGCT 2002  
1990 CGGCAGAGAAAGAGGCGGAGCGTCTCACTCTGAGGGTGAAGGCACTGTTTCAAGCTGCT 2049  
2003 CAACTATGAGCGGCAAAACATCTCTCACCTTATGGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACTACGAGCGGCGCGCGGCCCTCTTGGGCGCTCTGTGTGCTGGCGCTGAGCGA 2109  
2063 CATCTAAGAGACTGGCGGCGCTTGTGCTGCGTGTGCTGCTCTGACCCAGACACCCAG 2122  
2110 TATCCAGAGGCGCTGGCGACCTTTCGTGCTGCGTGTGCGGGCCAGGACCCGCGCTGA 2169  
2123 GATGTACTTTTGAAGCGAGATGTGACCGGGGCTATGATGCCATCCCGCCAGGGTAACT 2182  
2170 GCTGTACTTTTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTCAAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGAGGTCTATCGCAGCATCATCA---AACCCAGACACAGTACTGCGTGGTGGTA 2286  
2243 TGCAGTGTCCGGAGAGATAGCCAAAGGCCAAGTCCCAAGTCTCTTTAGGAGACAGGTCAC 2302  
2287 TCCGTGTGTCCAGAGGCGGCCCATGGGACGCTCCGCAAGGCTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAAGCATCTGCAGATTGAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTATATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGCATGTAGGACTCCGTTGTCATCGACGACAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCGACCGCTGAGGAGTGCCTGCTCATCGACAGAGTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCCTGTTTGACTCTTCTCTGCACTCTCTCGCTCACAGTGTCTTAAAGATTGGTGA 2482  
2461 CAGTGGCTCTTTCGAGCTCTTCTACGCTTCATGCGCACACGCGCTGGCGCATCAGGG 2520  
2483 CAGTGTCTATACGACGTCGAGGCGCATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCAGTCCAGGGGATCCCGCAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGGAGACATGAGAAACAAGCTCTTCTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCCTGTGCTAAGCGACATGAGAACAAAGCTGTTTTCGCGGATTCGCGGAGACGGCT 2640  
2603 GCTTTTACGTTTTGTTGATGATTTCTGTTGGTGAACGCTTCACTTGAGCCAAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTTGTGGTGACACCTCACTCACCACCGCAAAAC 2700  
2663 CTTCTCTCAGACCCCTGGTCCATGCGCTTCTGAGTATGGGTGATGATAAACTTGCAGAA 2722  
2701 CTTCTCTCAGGACCTTGTGTCGAGGTGCTCTGAGTATGGTGTGCTGAGTTCGGA 2760  
2723 GACGTGTGAACTTCTCTGTGAGCCTGTGTAACCTGGGTGTGAGCTCCATACAGCT 2782  
2761 GACGTGTGAACTTCTCTGTAGAAGACAGGCGCTTGGTGGCACCGCTTTTGTTCAGAT 2820  
2783 GCCTGCTCAGCTGCTTTTCTCTGTTGGTGTGCTTCTGCTGACACTCAGACTTTGGAGGT 2842  
2821 GCGGCGCCACGGCTTATTCCTCTGTTGGGCTGCTGCTGTGATACCGGACCTTGGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACTCAATTAAGACGAGCTCACCTTTCAGAG 2902  
2881 GCAGAGCAGTACTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGAGACCATCGGAAACAAGCTCTGCTGCGTCTTGGGTGAGAGTG 2962  
2941 CCGCTTCAAGGCTGGGAGGAACATGCGTGGCAAACTCTTTGGGGTCTTGGGCTGAAGTG 3000



2963 TCACGGTCTATTCTAGACTTCAGGTGAACAGCTCCAGACAGTCTGACATCATATATA 3022  
3001 TCACAGCCGTGTTCTGATTTCCAGTGAACAGCTCCAGACAGTGTGACCAACATCTA 3060  
3023 CAAGATCTTCTGCTTCAGGCTACAGGTTCCATGATGTGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTTCTGCTTCAGGCTACAGGTTTCAGCATGTGTGATGCTCCCATTTCA 3120  
3083 CCAGCGTGTAGGAGAGACTCAGATTTCTTCTGGCATCTCTCCAGCCAGCATCTG 3142  
3121 TAGCAAGTTTGAAGAACCCACATTTTCTGCGGTCTATCTGACAGGCTCCCT 3180  
3143 CTGCTATGCTATCTGATGCTCAAGATTCAGGATGACATAAAG-----GCCTC 3193  
3181 CTGCTACTCCATCTGAAGCGCAAGACGAGGATGCTGGGGGCAAGGCGCGC 3240  
3194 TGGCTCTTTCTCTGAAAGCGGCAATGGCTGTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTGCTCGAGGCGGTGAGTGGCTGTGCTGCTGCTCAAGCT 3300  
3254 GGTGCTCATTTCTGATCATCTCAAAATGCTCTCTGGGACTCTGAGGACAGCCAAAAC 3313  
3301 GACTGACACCGTGTACCTAGTGCCTCTCTGGGTCTCTGAGACAGCCAGACGA 3360  
3314 GGTGCTCGGAGCTCCAGAGGCGCAATGACCATCTTAAAGCTGAGTGAACCCAGC 3373  
3361 GCTGAGTCCGAAAGCTCCCGGGGACAGCTGACTGCTGCTGGAGGCGGAGCCAAACCCGGC 3420  
3374 CTTAAGCACAGACTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGACTGA 3454

AX133979 4015 bp DNA linear PAT 15-MAY-2001  
Sequence 1 from Patent EP1108789.  
AX133979  
AX133979.1 GI:14139919

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 Chang, S.Y. and Santini, C.D.  
Quantitation of htert mRNA expression  
Patent: EP 1108789-A 1 20-JUN-2001;  
F. HOFFMANN-LA ROCHE AG (CH)  
Location/Qualifiers  
1..4015  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
COUNT 663 a 1363 c 1275 g 714 t  
IN

Very Match 45.8%; Score 1600.8; DB 6; Length 4015;  
Est Local Similarity 68.9%; Pred. No. 0;  
Matches 379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGCGCTTTGAGCACAAATGACCCGGCTCTGTTGCCCCCGGGTGGGCTCTCTGCTGCG 82  
40 CCGGGCCACCCCGGATGCGCGGCTCCCGCTGCGAGCGCTGGCTCCCTGCTGCG 99  
83 CAGCGATACCGGAGGTGTGCGCTGGCAACCTTTGCGGGCGCTGGGGCCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGCGCTGGCAACCTTTGCGGGCGCTGGGGCCCGAGGG 159  
143 CAGGCGGCTTGTGCAACCCGGGACCCGAAAGATCTACCGCATTTTGTGTGCCAATGCT 202  
160 CTGGCGGCTGTGAGCGCGGGGACCCCGGCGCTTTCGCGCGCTGTGTGCGCGAGTGCCT 219

QY 203 AGTGTGATGACATGAGGCTTCAGAGCTTCACCTCGGACCTTTCTCTCCACAGGTGTC 262  
DB 220 GGTGTGCTGCTGCTTGGAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
QY 263 ATCCCTGAAGAGCTGTGTGCGGCTTGTGAGAGACTCTGAGAGCGCAACAGAGAGAA 322  
DB 280 CTGCTGAAGAGCTGTGTGCGCGCGAGTGTGAGAGCTGTGAGAGCGCGCGCGCG 339  
QY 323 GGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 382  
DB 340 GGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 399  
QY 383 CACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442  
DB 400 CACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459  
QY 443 TGTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502  
DB 460 GCGGTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519  
QY 503 CTGTGCTCTTTATCTTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562  
DB 520 CTGTGCTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579  
QY 563 GTACCAAAATTTGTGCCACACGAGATATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTG 622  
DB 580 GTACCAAGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639  
QY 623 CGACCCGTGGGAGGAAATTTCACTAATCTTACACAGATCAAGAGCAGTAG 682  
DB 640 GCGTCTGGGATGC-----GAAACGGGCTTGGAAACCATAG 672  
QY 683 TCGCCAGGAAGCACGAAACCCCTGGCTTCCCATCTCGAGGTACAAAGAGGCATCTGAG 742  
DB 673 CGTCAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732  
QY 743 TCTCACAGTACAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 802  
DB 733 CAGTCCAGGCGGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792  
QY 803 GGAGGAGGACCCACAGGAGGTGTACCAACCCCATCAGGCAATCATGGGTGGCGCAAG 862  
DB 793 GGAGGAGGACCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 852  
QY 863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
DB 853 TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
QY 914 AGGAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
DB 913 GGTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972  
QY 971 CTCACAT---CTCTGCTGTACACACCCCGCAAAATGCTTTCAGCTCAGGCCATTTAT 1027  
DB 973 CCCCCCATCCACATCGCGGCGCACCATGCTCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
QY 1028 TGAGCCAGACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1087  
DB 1033 CGAGCCAGCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1089  
QY 1088 CTTACTCAGCAACCTCCAGCTTAATTTGACTGGGGCCAGGAGACTGCTGGAGATCATCTT 1147  
DB 1090 CTTACTCAGCTCTCTGAGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
QY 1148 TCTGGCTCAAGGCTTAGGACATCAGGACCATCTGAGGACACACGCTTATCGGCTG 1207  
DB 1150 TCTGGGTTCAGGCGCTGATGCGAGGACTTCCCGCGGTTCGCCCGCTGCGCCAGCG 1209  
QY 1208 ATACTGGCAGATGCGGCGCTGCTTCCAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
DB 1210 CTACTGGCAATGCGGCGCTGCTTCTGAGCTGCTTGGGAACCAAGCAGCTGCTGCTGCT 1269  
QY 1268 TGTGAGACTCTCAGGTCAATTTGAGGTTTCGAAAGCAACCAAGAGGTGACATGCTC 1327

[illegible]

:ch completed: February 2, 2004, 01:50:58  
time : 12479 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: February 1, 2004, 19:29:41 ; Search time 863 Seconds  
(without alignments)  
10935.383 Million cell updates/sec

le: US-09-042-460-1

fect score: 3496

ence: 1 GAATTCGGGCTGGAGGCC.....CGAGCTCGTACCAAGCTT 3496

ring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

ched: 2552756 seqs, 1349719017 residues

al number of hits satisfying chosen parameters: 5105512

imum DB seq length: 0

imum DB seq length: 2000000000

-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : N Geneseq\_19Jun03.\*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1600.8	45.8	3955	19 AAV22379	Human telomerase r
2	1600.8	45.8	4015	20 AAV230154	cDNA encoding a hu
3	1600.8	45.8	4015	20 AAZ08150	Human telomerase r
4	1600.8	45.8	4015	20 AAZ20279	Human telomerase r
5	1600.8	45.8	4015	20 AAZ00724	Human telomerase c
6	1600.8	45.8	4015	22 AA45901	Human hTERT gene
7	1600.8	45.8	4015	24 AAD46821	Human telomerase r
8	1600.8	45.8	4015	24 ABA97534	Cancer cell discr

9	1600.8	45.8	4015	25 ABZ22474	Human telomerase r
10	1600.8	45.8	4015	25 ABZ18391	Group III cDNA can
11	1600.8	45.8	4023	19 AAV60320	Human telomerase g
12	1600.8	45.8	4027	20 AA89424	Human EST coding
13	1600.8	45.8	4027	21 AA829388	hEST2, a human tel
14	1600.8	45.8	4042	20 AAV72117	Human catalytic te
15	1600.8	45.8	4070	24 ABL53711	Human telomerase c
16	1599.2	45.7	3798	19 AAV27876	Human telomerase p
17	1596	45.7	3399	24 ABZ35720	Human telomerase r
18	1596	45.7	3399	24 ABX09963	Human telomerase r
19	1596	45.7	3399	24 ABV78144	Human telomerase r
20	1596	45.7	3399	24 ABL91685	Human polynucleoti
21	1596	45.7	3964	20 AAX18254	Human telomerase c
22	1594.6	45.6	3396	20 AAX18266	Telomerase coding
23	1594.6	45.6	3396	22 AAH44366	Human telomerase n
24	1594.6	45.6	3396	22 AAH48235	Heart muscle cell
25	1594.6	45.6	3396	22 AAH49601	Human coding seque
26	1594.4	45.6	4037	19 AAV22428	Human telomerase r
27	1593.6	45.6	3453	25 ABZ76217	Human TERT coding
28	1593.6	45.6	13766	24 AAD46790	pGRN145 plasmid DN
29	1584.8	45.3	8742	24 AAD46793	PWGB5a plasmid DNA
30	1526.8	43.7	3918	20 AAX18269	Telomerase coding
31	1526.8	43.7	3918	20 AAX18278	Telomerase coding
32	1508.8	42.4	3203	20 AAX18268	Altered C-terminus
33	1480.6	42.4	3500	20 AAX18275	Telomerase (ver. 2
34	1462.6	41.8	7498	21 AAB63786	Nucleotide sequenc
35	1462.6	41.8	7498	24 ABL50133	Mouse telomerase g
36	1438	41.1	3167	20 AAX18271	Altered C-terminus
37	1438	41.1	3167	20 AAX18280	Altered C-terminus
38	1377.2	39.4	3323	20 AAX18277	Telomerase coding
39	1374.4	39.3	7688	20 AAX18351	Human telomerase r
40	1323.2	37.8	3855	19 AAV22382	Telomerase coding
41	1312.4	37.5	7797	20 AAX18350	Truncated telomera
42	1297.8	37.1	3069	20 AAX18276	Truncated telomera
43	1297.8	37.1	3069	20 AAX18267	Human telomerase p
44	1264.6	36.2	2848	19 AAV27872	Truncated telomera
45	1230.2	35.2	3033	20 AAX18270	Truncated telomera

#### ALIGNMENTS

#### RESULT 1

AAV22379  
ID AAV22379 standard; cDNA; 3955 BP.

XX AAV22379;

XX AC

XX 13-AUG-1998 (first entry)

XX Human telomerase reverse transcriptase encoding cDNA.

DE Human; telomerase reverse transcriptase; hTERT; TERT; diagnosis;

KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 56..3454

XX /\*tag= a

XX /\*product= "telomerase reverse transcriptase"

XX GB2317891-A.

XX 08-APR-1998.

XX 01-OCT-1997;

XX 97GB-0020890.

XX 97US-0915503.

XX 96US-0724643.

XX 97US-0844419.

XX 97US-0846017.

XX 97US-0851843.

400 CACCACGCGTGCAGCTACCTGCCCAACACGGTGACCGACGCACTGCCGGGAGCGG 459

443 TGACGATGCTACTGTGTTGAGCCGAGTGGGCGACGACCTGCTGGTCTACCTCTGCTGCACA 502

460 GGGTGGGGGCTGCTGCTCGCGCGGTGGGCGACGAGTGTGGTTCACCTCTGCTGCACG 519

503 CTGTGCTCTTTATCTCTCTGGTCCCGCCAGCTGTGCCTACACAGTGTGTGGGTCTCCCT 562

520 CTGCGGCTCTTTGCTGGTCCCTCCAGCTGCCTACACAGTGTGGGCGCGCGCT 579

563 GTACCAAAATTTGTCACACACGATATCTGGCCCTCTGTTCGCCTAGTTTACAGGCCAC 622

580 GTACCAAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCGAA 639

623 CCGACCCGTGGGACGAATTTCTACTAAGCTTAGTCTTTCAACACAGATCAAGAGCAGTAG 682

640 GCGTCTGGGATG-----GAACGGGCTTGGAAACCATAG 672

683 TCGCCAGGAACACCAAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 742

673 CGTCAGGAGGCCGGGGTCCCCCTGGCGTCGACGCCCGGGTTCGAGGAGGGCGCGGGG 732

743 TCTCACCACTACAAGTGTGCTTCAGCTAAGAAGCCAGATGCTATCTGTCCGAGAGT 802

733 CAGTGCCAGCCGAAGCTCCCGTTGCCCAAGAGGCCAGGCGTGGCGCTGCCCTCAGCC 792

803 GGAGAGGGACCCCAACAGCAGGTGCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862

793 GGAGCGGACCCCGTTGGGACGGGTCTTGGGCCCAACCCGGGCGAGGACGGCTGGACCGAG 852

863 TCCTGCTCGGTCCCGCCAGGTGCCTACT-----GCAGAGAAAGATTTGCTTCTAA 913

853 TGACCGTGGTTTCTGTGTGGTGTCACTGCCACAGACCCGCCGAAGAACCACTCTTTGGA 912

914 AGGAAAGGTCCTGACCTGAGTCTCTC---TGGGTCGTTGCTGTATAACACAAAGCCACG 970

913 GGGTGGCTCTCTGGACCGGCGCACTCCACCCATCTCGTGGCGCCAGACACACAGCGGG 972

971 CTCACAT---CTCTGTGTCAACCCCGCCAAATGCCCTTCAGTCTAGGCCATTTAT 1027

973 CCCCCATCCACATCGCGGCCACACGCTCCCTGGGACACGCTGTCTCCCGGTGTACG 1032

1028 TGAGACCAGACATTTCTTTACTTCCAGGGGAGATGGCCAAAGAGGCTCTAAACCCCTCAT 1087

1033 CGAGACCAAGCATTCTCTTACTCTCTAGCGA---CAAGAGCAGCTGGCGGCTCTCTT 1089

1088 CCTACTCAGCAACCTCCAGCCTAATCTGACTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147

1090 CCTACTCAGCTCTCTGAGGCCACGCTGACTGGCGCTCGGAGGCTCTGTGGAGACCATCTT 1149

1148 TCTGGGCTCAGGCTTAGGACTCAGGACCACTCTGAGGACACACGCTCTATCGCGTGG 1207

1150 TCTGGGTTCCAGGCCCTGGATGCGAGGACTCCCGCAGGTGGCCCGCTCTGCCCGAGG 1209

1208 ATACTGGCAGATCGGCGCCCTGTTCAAACAGTGTGTGGTAAACCATGACAGATGCGCAATA 1267

1210 CTACTGGCAATGGGCGCCCTGTTTCTGGAGTGTGTTGGGAACCAACGGCGAGTGGCCCTA 1269

1268 TGTCAAGCTCTCAGGTCAATTCAGGTTTCGAAACAGCAAAACCAAGGTGACAGATGC 1327

1270 CGGGGTGCTCTCAAGAGGCACTGCCGCTGCGAGCTGGGTACCCCGAGCGCGGTGT 1329

1328 CT-----TGAAACACAGGCC 1342

1330 CTGTGCCCGGAGAACGCCAGGGCTCTGTGGCGGCCCGCCGAGGAGGAGACACAGACC 1389

1343 ACCGACACCTCATGATTTGCTCCGCTGACAGCAGTCTCCCTGGCAGGTATATGGTTTCT 1402

1390 CCGTCCCTGGTGCAGCTGCTCCGCCACACACAGCAGCCCTGGCAGGTGTACGGCTTCGT 1449

1403 TCGGCCCTGTCTCTGCAAGGTGGTGTCTGTAGTCTCTGGGGTACACAGGCAATGAGCG 1462

1450 GGGGCTCTGCTGGCGCGGTGTGGCCCCCAGGCTCTTGGGGCTCCAGGCAACAGAAAC 1509

1463 CCGCTCTTTAAGAACTTAAGAAAGTTTCACTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAAACCAAGAAAGTTTCACTCTGTTGGGGAATACGGCAAGCTATCACT 1569  
1523 GCAGGAAGTATGTTGAAGATGAAGTATGAGGATGCGCTCCGAGCAGAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTCCGAGAGCCGAGG 1629  
1583 GAAGGACCGTGTCCCGCTGCAGAGACCGTGTGAGGAGAGAGTCTGGTCACTGCTTCC 1642  
1630 GGTGGCTGTGTTCCGCGCCAGAGACCGTGTGCGTGAGGAGATCTTGGCCAAAGTTCC 1689  
1643 GTTCTGCTGATGGAACACATACCTGTGATGCTTACCTGCTTACCTTTTATCATCAGAGA 1702  
1690 GCACCTGCTGATGAGTGTGATCTGTGCTGAGCTGTCTCAGGTCTTCTTTTATGTCAGGGA 1749  
1703 GAGCACATTCAGAGAAACAGGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGGAGTCAGGCAACACCTTCAGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGGAATCAGACAGCATTTGAAGAGGTGACGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTGAGCATACACAGGACACTGGTACGCTACGCTACGCTACGCTACGCTACGCTACGCT 1882  
1870 GGTGAGGAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1929  
1883 CAGGCCAACGGGCTCGGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942  
1930 CAGGCTGAGCGGCTCGGCCGATTTGAACTGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1989  
1943 GGGCAGAGGAAGCAGGCGCCAGCATTTCAACCAGCGCTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACTCGAGGGTGAAGGCACTGTTTCAGCGTCT 2049  
2003 CACTATGACGAGCAAAACATCTCACTTATGAGGCTCTTCTGCTGCTGCTGCTGCTGCTGCT 2062  
2050 CAACTACGAGCGGCGGCGCGGCGGCTCTGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2109  
2063 CATCTACAGACTGCGGCGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCACAGGCTGCGGCGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATGTACTTTTAAAGCAGATGTGACGCGGCTTATGATGCTATCCCTCCGAGGTAAGCT 2182  
2170 GCTGTACTTTTCAAGTGTGATGTGACGCGGCTTACGACACCATCTCCCGAGGAGCT 2229  
2183 GGTGGAGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCGAGTA 2242  
2230 CAGGAGGCTCATCGCAGCATATCA---AACCCAGAAACAGTACTGTGCTGCTGCTGCTGCT 2286  
2243 TGCAAGTGGTCCGAGAGATAGCCAGGCGCAAGTCCCAAGTCTTTTAGGAGCAGGTCA 2302  
2287 TGCGGTGCTCAGAGGCGCGCTATGGGCACTCCGCAAGGCTTTCAAGAGCCAGCTCTC 2346  
2303 CAGCTCTGCTGCTCAGGCTATGAGGCGAGTCTTACGATCTGCGAGGATTCAGA 2362  
2347 TACTTTCAGAGACTTCAGCGCTACATGAGCAGTTCGTGGCTCACCTGCGAG-----GA 2400  
2363 TGCCAGTGCATCAGGAATCCCTGCTGCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGAGTCCGCTGCTGCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGGCTGTTGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2482  
2461 CAGTGGCTCTTTCAGAGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGGTGCTATACGAGTGCAGGCTATCCCGAGGCTCCAGCTATCCAGCTCTGCTGCTGCTGCT 2542  
2521 CAGTCTAGTCCAGTGCAGGAGTCCCGAGGCTCCATCTCTCTCAGCGCTGCTGCTGCTGCT 2580  
2543 CAGTCTGTTTCGAGACATGAGAGAACAAAGTGTGTTGCTGAGGTGCGAGCGGAGTGGGTT 2602

2581 CAGCTGTGCTACGCGCATGAGAACAAAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGTGATGACTTTCTGTTGGTGAAGCTCCTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTGTTGGTGAACCTCACCTCACCCACGCGAAAC 2700  
2663 CTTCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2722  
2701 CTTCTCAGAACCTTGGTCCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2723 GACAGTGTGAATCTTCCCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2782  
2761 GACAGTGTGAATCTTCCCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCT 2842  
2821 GCGGCGCCACGCTTATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGCT 2902  
2881 GCAGAGCGACTACTCCAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
2903 TGTCTTCAAAGCTGGGAGAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGCTGGGAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000  
2963 TCACGCTCTATTTCTGAGACTTGCAGGTGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3022  
3001 TCACAGCTGTTTCTGAGTTTGCAGGTGAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3060  
3023 CAAAGTCTTCTGCT 3082  
3061 CAAAGTCTTCTGCT 3120  
3083 CCAGGCTGTAGGAGAACTCAGATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
3121 TCAGCAAGTTTGAAGAAACCCCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
3143 CTGCTATGTATCTTGAAGGCTCAAGATCCAGGATGACACTAAAG-----GCTC 3193  
3181 CTGCTACTCATCTTGAAGCAAGAACGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3240  
3194 TGGCTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
3241 CGGCTCTGCT 3300  
3254 GGTGCTCATCTGCT 3313  
3301 GACTGACACCGCT 3360  
3314 GCT 3373  
3361 GCTGAGTGGAAAGCTCCCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CCTAAGCAGAGCTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3454

## RESULT 2

AAZ30154

ID AAZ30154 standard; cDNA; 4015 BP.

XX AAZ30154;

AC AAZ30154;

XX 26-JAN-2000 (first entry)

DT 26-JAN-2000 (first entry)

XX cDNA encoding a human telomerase reverse transcriptase (TRT).

DE Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

XX Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;

ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 56..3454/\*tag= a  
/product= "telomerase reverse transcriptase"

W09950392-A1.

07-OCT-1999.

30-MAR-1999; 99WO-US06898.

31-MAR-1998; 98US-0112006.

(GERO-) GERON CORP.

Gaeta FCA;

WPI; 1999-610845/52.

P-PSDB; AAY43621.

Eliciting an in vivo immune response for prevention and treatment of cancers

Disclosure; Fig 2; 26pp; English.

The present sequence encodes a human telomerase reverse transcriptase (TRT) polypeptide. The protein is used in the method of the invention. The specification describes a method for activating a T lymphocyte, comprising contacting the T lymphocyte with a dendritic cell that expresses a TRT peptide in the context of a MHC class I or MHC class II molecule. The protein causes induction of an in vivo immunological response to telomerase activity. Cancer cells are characterized by expression of endogenous TRT gene and the presence of detectable telomerase activity. Therefore, by eliciting a specific immune response to TRT or to TRT-expressing cells, it is possible to selectively target proliferating cells for immunological destruction. The method is used for eliciting an in vivo immune response to telomerase by activating a T lymphocyte, and is useful for prevention and treatment of cancers and other proliferation diseases/conditions.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGTAGCACAATGACCCGCGCTCTCGTTGTCGCCCGCGTGCCTCTCTGTGCG 82  
|||||  
40 CCGGCGCACCCCGCGATGCGCGGCTCCCGCTGCGGACCGTGCCTCCCTGTGCG 99  
|||||  
83 CAGCGGATACCGGAGGTGTGCGCGCTGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 142  
|||||  
100 CAGCCNACTACCGAGGTGCTGCCCTGSCACAGTTCTGTGCGGCGCTTGGGCCCCAGGG 159  
|||||  
143 CAGGCGGCTTGTGCAACCCGGGACCCGAAGATCTACCGCACTTTGTGTTGCCAATGCTT 202  
|||||  
160 CTGCGCGCTGTGTGACGCGCGGGAACCGCGCGCTTTCGCGCGCTGTGTGCGCCAGTGCCT 219  
|||||  
203 AGTGTGCATGCTAGGCTGCTACAGCTTCCAGCTTCCAGCTTTCCTTCCACAGGTTGC 262  
|||||  
220 GGTGTGCTGCTTGGAGCGACGCGCGCGCCCGCGCGCCCTCTCTTCCGCGAGGTGC 279  
|||||  
263 ATCCCTGAAGAGCTGGTGGCCAGAGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAAA 322  
|||||  
280 CTGCTGTGAAGAGCTGGTGGCCCGAGTGTCTGAGAGGCTGTGCGAGCGCGCGCGAGAAA 339  
|||||  
323 CGTGTGCGCTTTTGGCTTTGAGCTCTTAACGAGCCAGAGCGCGGCTCTCCCATGGCTT 382  
|||||  
340 CGTGTGCGCTTTGCGCTTCTGCTGCTGAGAGGGGCGCGCGGGGCGCCCGCGAGGCTT 399  
|||||

QY 383 CACTAGTAGCGTGTAGTACTTGGCCCAACTGTATTGTAGACCCCTGGCTGTCTCAGTGG 442  
|||  
Db 400 CACCAACACGCGTGTGCGAGTACTTGGCCCAACACGCTGACCGACGCTGCGGGGAGCGG 459  
|||  
QY 443 TGATGATGCTACTGTGTAGCGAGTGGCGGACGACCTGTGTGTCTACCTGCTGTGCAACA 502  
|||  
Db 460 GCGTGGGGCTGCTGTGCGCGCGTGGCGGACGAGTGTGTGCTTACCTGTGTGCGCAG 519  
|||  
QY 503 CTGTGCTCTTTTATCTTCTGTGTGCGCCCGCAGCTGTGCTTACGAGGTGTGTGGTCTCCCT 562  
|||  
Db 520 CTGCGCGCTCTTTGTGTGTGCTGCCCTCCAGCTGTGCTTACCGAGGTGTGCGGCGCCCGCT 579  
|||  
QY 563 GTACCAAAATTTGTGCCACCAACGATATCTGTGCTCCGCTAGTTTACAGGCCCCAC 622  
|||  
Db 580 GTACCACTGCGGCTGTGCACTCAGGCCCGGCCCGCCACACGCTAGTGGACCCCGGAG 639  
|||  
QY 623 CCGACCGTGGGAGGAAATTTCACTAACCTTAGTGTCTTACCAACAGATCAAGAGCAGTAG 682  
|||  
Db 640 GCGTCTGGGATGC-----GAAACGGGCTGTGAACCATAG 672  
|||  
QY 683 TCGCAGGAGACCCGAAACCCCTTGGCTTGCATCTCGAGGTACAAAGAGCATCTCTAG 742  
|||  
Db 673 GGTGAGGAGCGGGGTCCCTTGGCTTGCAGCCCGGGTGGAGAGCGCGGGG 732  
|||  
QY 743 TCTCACCAGTACAAAGTGTGCTTCACTAAGAGGCCAGATGCTATCTCTTCCGAGAGT 802  
|||  
Db 733 CAGTCCAGCGGAAGTCTGCGCTTGCACAGAGGCCAGCGTGGCGCTGCGCTGAGCC 792  
|||  
QY 803 GGAGGAGGAGCCACAGGCGAGTGTACCAACCCCATCAGGCAATCATGTGGTGTGCAAG 862  
|||  
Db 793 GGAGGAGGAGCCCGTGGGAGGGGTCTTGGGCGCCACCCCGGAGAGCGGTGGAGCCGAG 852  
|||  
QY 863 TCTGTGCTGCTCCCGGAGTGTGCTACT-----GCAGAGAAAGATTTCTTCTTAA 913  
|||  
Db 853 TGACCGTGGTTCTGTGTGTGTGCTACCTGCGAGACCCGCGGAAGCCACCTCTTTGA 912  
|||  
QY 914 AGAAAGGTGTGCTGACCTGAGTCTCTC-----TGGGTGGTGTGCTGTAAACAAGCCAG 970  
|||  
Db 913 GGGTGTGCTCTGTGCAACGCGCCACTCCACCCATCCGTGGGCGCGCCAGCACCAACGCGGG 972  
|||  
QY 971 CTCCACAT---CTCTGTGTGTGACACCCCGCCCAAAATGCTTTTCACTCAGGCTCAGGCCATTTAT 1027  
|||  
Db 973 CCCCCATCCACATCGCGGCCACCACTCCCTGGGACACGCTTGTCTCCCGCTGTGTACGC 1032  
|||  
QY 1028 TGAGACGAGATTTCTTTATCTTCCAGGGAGATGCGCAAGAGCGTCTAAACCCCTCAT 1087  
|||  
Db 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
|||  
QY 1088 CCTACTCAGCAACCTCCAGCTTAACCTGCTGGGCGCAGGAGCTGGTGGAGATCATCTT 1147  
|||  
Db 1090 CTTACTCAGCTCTCTGAGGCGCCAGCTGACTGTGCGCTCGAGGCTCTGTGAGACCATCTT 1149  
|||  
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGAACCATCTGCGAGGACACACCGCTTATCGGCTG 1207  
|||  
Db 1150 TCTGGTTCCAGGCGCTGTGATGCCAGGACTCCCGCGAGTTGCCCGCTGCGCCAGCG 1209  
|||  
QY 1208 ATACTGAGATGCGGCGCCCTGTTCACACAGCTGCTGTGAACCATGAGAGTGGCAATA 1267  
|||  
Db 1210 CTACTGCAAAATGCGGCGCTCTTCTGTGAGCTGTCTGTGGAACCCAGCGCAGTGGCCCTTA 1269  
|||  
QY 1268 TGTGAGACTCTCTCAGGTTCACATTTGCAAGTTCGAGTTTCGAACAGCAAAACCAACAGGTGACAGATGC 1327  
|||  
Db 1270 CGGGGTGTCTCTCAAGACGCACTGCGCGCTGCGAGTGTGCGTCAACCCAGAGCGGCTGT 1329  
|||  
QY 1328 CT-----TGAAACACAGGCC 1342  
|||  
Db 1330 CTGTGCGCGGAGAGCCCGAGGGCTCTGTGGCGGCGCCCGAGAGGAGGAGACACAGACC 1389  
|||  
QY 1343 ACCGACCTCATGATTTGCTCCGCTGTGACAGAGTCTCCCTGGCAGGTATATGTTTTCT 1402  
|||  
Db 1390 CCGTGTGCTGTGTGAGCTGTCTGCGCCAGCAGACAGACCCCTGTGCGAGTGTACGGCTTCT 1449  
|||

1403 TCGGGCTGTCTCTCAAGGTGCTGTCTGTAGTCTCTGGGTACCAAGCACATGAGCG 1462  
1450 GCGGGCTGTCTCTCGCGGGTGTGTCCGCCAGAGCTCTGGGGCTTCCAGGCACACGACG 1509  
1463 CGGCTTCTTTAAGACTTAAGAAAGTTTCATCTCGTGTGGGAAATACGGCAAGCTTACCT 1522  
1510 CGGCTTCTCTCAGGAACACCAAGAAAGTTTCATCTCCCTGGGAAAGCATGCCAAGCTCTCGCT 1569  
1523 CGAAGAACTGATGTGGAAGATGAAGATGAGAGTATGCCACTGGCTCCGACGACGCCGGG 1582  
1570 GAGAGAGTACGCTGGAAGATGAGCGTCCGGAGCTGCGCTTGGCTGCGAGAGCCGAGG 1629  
1583 GAAGGACCGTGTCCCGCTGTCAGAGACACCGTCTGAGGAGAGGATCTCTGGCTACGTTTCT 1642  
1630 GGTGTGCTGTGTCTCCGCCGACAGACCGTCTGCGTGAGGAGATCTTGCCCAAGTTTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGTGATACGCTGCTTAGGTTCATCTTTTACATCAGGA 1702  
1690 GCACTGGCTGATGAGTGTGTAGCTGTCTGAGTCTGCTCAGGTCTTTTCTTTTATGTACGGA 1749  
1703 GAGCACATTTCCAGAAAGACAGGCTCTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTTTCAAGAGACAGGCTCTTTTCTTACCGNAGAGTGTGTGGAGCAAGTTGCA 1809  
1763 GAGCAATGAGAGTACGCAACACCTTGAAGAGTGTGGCTACGGAGCTGTCAAGAGGA 1822  
1810 AAGCAATGGAATCAGACAGCACTTGAAGAGGAGTGTGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGCATCACAGGACACCTGTGATAGCATGCCATCTGCAAGTGTGGCTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGAGAGCAGAGCCGCGCTGTGACGTCCAGATCTCGCTTCATCCC 1929  
1883 CAAGCCCAACCGGCTCGCGCCCATTTGGAACATGAGTATAGCATGGGTACCGAGCTTT 1942  
1930 CAAGCTCGACGGGCTCGCGCCGATTTGGAACATGAGTGTGCTGCGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGCAGGCGCCAGCATTTCAAGGCTCTCAAGACTCTCTTCAAGATGCT 2002  
1990 CGCAGAGAAAGGGCGGAGCGTCTCACCTCAGAGGTGAAGCACTGTTCAGCGGTCT 2049  
2003 CAACATAGCGGAGCAAAACATCTCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACATAGCGGCGGCGGCGCGGCTCTCTGGGCGCTCTGTGTGTGGCTTGACGA 2109  
2063 CATCTACAGGACCTGGGCGGCTTTGTGCTCGGTGTGGGTGTCTGTGAACAGACACCCAG 2122  
2110 TATCCACAGGGCTCGGCGACCTTGTGCTGCTGCTGTGCGGGCCAGGACCGCGGCTGA 2169  
2123 GATGTACTTTTGTAAAGCAGATGTGACCGGGCTTATGATGCGCATCCCCAGGTAAGCT 2182  
2170 GCTGTACTTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCAGACAGGCT 2229  
2183 GGTGAGGTTTGTGCAATATGATCAGGCACTCGGAGAGCAGTGTATCTGCCCGAGTA 2242  
2230 CACGGAGGTTCATCGGCAGCATCATCA--AAACCCAGAAACAGTACTCGTGTGCGTGA 2286  
2243 TGCAGTGTGCGGAGATAGCCAGGCGCAAGTCCAGATGCTTTAGGAGACAGGTCTAC 2302  
2287 TGCCGTGTGCAAGAGCGGCCCTAGGCGAGTCCGCAAGGCTTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATATAGTGGGCGAGTCTCTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATCGACAGTTCGTGCTCACCTGCGAG--GA 2400  
2363 TGCCAGTGCATGAGGACTCGGTGTGATCGAGCAGAGCATCTCTATGATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGGATGCGGTGTCTATCGAGCAGAGTCTCTCCCTGATGAGCCAG 2460  
2423 CAGCAGCTTTTGAATCTTCTTCTGCACTTCTCGTCAAGTGTCTGTAAGATGAGTGA 2482  
2461 CAGTGGCTCTTCGAGGCTTCTTCTTACGCTTTCATGTGCCACCAAGCGGTGCGCATCAGGG 2520  
2483 CAGGTGTCTATAGCGAGTGCAGGGCATCCCCAGGGCTCCAGCGCTATCCACCTGTCTGTG 2542

Db 2521 CAAGTCTTACGTCTCAGTGCACGGGATCCCGAGGGCTCCATCTCTCCACGCTGTCTTG 2580  
QY 2543 CAGTCTGTGTTTCGAGACATGAGAACAAAGCTGTTTGTCTAGGTGACGCGGATGGTT 2602  
Db 2581 CAGCTGTGCTTACGCGACATGAGAACAAAGCTGTTTTCGGGATTCGGCGGACGGGT 2640  
QY 2603 GCTTTTACGTTTGTGATGACTTCTGTTGGTGAAGCGCTCACTTGGACCAAGCAAAAC 2662  
Db 2641 GCTCTGTGTTTGTGTGATGATTTCTTGTGGTGAACACCTCACCTCACCCACGCAAAAC 2700  
QY 2663 CTTCCTCAGCACCCCTGTCTCATGGCTTCTGAGTATGGGTGATGATATAAATCTTGAGAA 2722  
Db 2701 CTTCCTCAGGACCCCTGTCTCGAGGTGTCCCTGAGTATGCTGGTGTGAACTTGGGAA 2760  
QY 2723 GACAGTGTGAACTTCTCCTGTGAGCGCTGTACCTTGGGTGAGCTCCATACAGCT 2782  
Db 2761 GACAGTGTGAACTTCTCCTGTAGAAAGCAGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGAT 2820  
QY 2783 GCCTGTCACTGCTGTTTCCCTGGTGGCTTGTCTGTGGACACTCAGACTTTGGAGGT 2842  
Db 2821 GCCGCGCCACGGGCTATTCTCCTGGTGGGCTGTCTGTGGATACCCGGAACCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACCTTCAGAG 2902  
Db 2881 GCAGAGGAGTACTTCCAGTATATGCCGGAACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAGCTGGGAGAACCATGCGGAACAGCTCCTCGGTCTTGGGTGAGGTG 2962  
Db 2941 CGGCTTCAAGCTGGGAGAACATGCGTGCAGAACCTTTCGGGTCTTGGGTGAGGTG 3000  
QY 2963 TCACGGTCTATTCTAGACTTGCAGGTGAACAGCTCCAGACAGTGTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGATTTGCAAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAGATCTTCTGCTTTCAGGCTTACAGGTTCATGATGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAGATCTCTGCTGTGAGCGGTACAGGTTTTCAGCATGTGTGCTGCTGCTCCCATTTCA 3120  
QY 3083 CCAGCGTGTAGGAAGAACCTCACATTTCTTCTGGGCATCATCTCCAGCCAAAGCATCTTG 3142  
Db 3121 TCAGCAAGTTTGAAGAAACCCACATTTTCTTGGCGGTCACTCTGTGACACGCGCTCCCT 3180  
QY 3143 CTGTATGTCTATCTTGAAGTCAAGATCCAGAAATGACATAAG-----GCCTC 3193  
Db 3181 CTGTACTCTCATCTGAAAGCAAGAACGAGGATGTGCTGGGGCCCAAGGGCGCCGC 3240  
QY 3194 TGGCTCTTCTCTTCAAGCCGACATTTGGCTCTCTACAGGCTTCTCTGCTCAAGCT 3253  
Db 3241 CGGCGCTCTGCGCTCGAGGCCGTGAGTGGCTGTGCTGCCACCAAGCATTCCTGCTCAGCT 3300  
QY 3254 GGCTGTCTATCTGTCTATCTAATAATGTCTCTGGACCTCTGAGGACAGCCCAAAAT 3313  
Db 3301 GACTCGACACCGTGTCACTTACCTTGTGCGCTTCTTGGGTCCTCAGGACAGCCAGACGA 3360  
QY 3314 GCTGTGCGGAGTCTCCAGAGGGGCAATGACCATCTTAAAGTGTGAGCTGACCCAG 3373  
Db 3361 GCTGAGTGGAGTCTCCGGGAGCAGCTGATGCTGCCCTGGAGCGCGGACCAACCCGCG 3420  
QY 3374 CCTAAGCACAGACTTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

## RESULT 3

AAZ08150

ID AAZ08150 standard; cDNA; 4015 BP.

XX AAZ08150;

AC AAZ08150;

DT 17-JAN-2000 (first entry)

XX Human telomerase reverse transcriptase cDNA.

DE



Human telomerase reverse transcriptase, hTERT; telomerase; hEST2; catalytic protein component; cell proliferative capacity; cell immortality; neoplastic phenotype; diagnostic application; prognostic application; telomerase related condition; cancer; therapeutic agent; telomerase expression; telomerase activity; ds.

**Homo sapiens.**

Location/Qualifiers

56.3454

/\*tag=

```

/product= "Human telomerase reverse transcriptase"
/transl_except=(pos:1877..1879, aa:Gln)

```

· WO9950279-A1.

07-OCT-1999.

31-MAR-1999: 99WO-US07160.

37-MAR-1998: 98US-0052919.

(GERO-) GERON CORP.

(DYTE-) UNIV TECHNOLOGY CORP.

Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB; Andrews WH;

WPI: 1999-610834/52.

P-PSDB: AAY28881.

Antisense polynucleotides for human telomerase reverse transcriptase used for diagnosing or treating cancer -

Claim 1: Fig 1: 31pp: English.

The present sequence encodes for human telomerase reverse transcriptase (hTERT), this catalytic protein component of telomerase and is also linked to the cellular lifespan. hTERT correlates with cell proliferative capacity, cell immortality and the development of a neoplastic phenotype. Human telomerase is a ribonucleoprotein complex composed of an RNA subunit and a protein subunit. Antisense oligonucleotides are useful for diagnostic or prognostic applications to telomerase related conditions, including cancer. They are also useful as therapeutic agents, for inhibition of telomerase expression and activity.

Sequence 4015 BP; 663 A; 1364 C; 1274 G; 714 T; 0 other;

Query Match 45.8%: Score 1600.8: DB 20: Length 4015;

```

{very MacCn
}est Local similarity 68.9%:
Pred. No. 0:
}

```

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGGCCCTTGAGCACAA TGACCCGGGCTCCTCGTTGCCCGGCTCTCTGCTGG 82

40 CCGGGCCACCCCGGATCCCGGGGATCCCCGCTGCCGAGCCGTGGCTCCCTGCTGG 99

[illegible]

100 GAGGCGCCCTGGTGTGCATCCTGGGCGTTGGGCGTCACGG  
15

20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050

23

[illegible][illegible]

100

.....

.....

340 CGTGCTGGGCTTCGGGTTCGGGCTGCTGGACGGGGCCCGGGGGGCCCCCGAGGCGCTT 399

383 CACTAGTAGGGTGGCTAGTACTCTGCCCCAACACTGTATTATGACACCCCTCGGTGTCAAGTGG 442

400 CACACCAAGGTGGCGACGCTACTCTGCCCAACAACGGTGACCGACACTCGGGGGAGCGG 459

443 TGCATGSAATGCTACTGTATTGAGCCGAGTGGGCGACGACTCTCTGGTCTACTCTGCTGGCAC 502

460 GGCCTGGGGGCTGCTGCTCGCGCGGTGGGCGACGAGTGTGGTATTCACTGCTGGCACG 519

503 CTGTGCTCTTTATCTTCCTGGTGCCTCCCGAGCTGTGCTTACAGAGTGTGGGTCTCCCT 562

520 CTGGCGGCTCTTTGTGCTGGTGGCTCCCACTGGCTTACAGGTGTGGGCGCGGCT 579

563 GTACCAAAATTTGTGCCACACGGAATATCTGGCCCTCTGTGTCCGTAGTTTACAGGCCAC 622

580 GTACCACTCGGCGCTGCCACTCAGCCCGGCCCGCCACACGCTAGTGGACCCCGAAG 639

623 CCGACCCGTGGCGAGGAATTCATTAACTTAGTTCTTACACAGATCAAGAGCAGTAG 682

640 GCGTCTGGGATGC-----GACCGGCCCTGGAAACCATAG 672

683 TCGCACGAAGACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 742

673 CGTCAGGAGCGCGGGTCCCCCTGGGCTGCCAGCCCGGTCGAGGAGGCGCGGGG 732

743 TCTCAACGATCAAGTGTGGCTTCAGCTTAAGAGCCAGATGCTACTCTCTCCGAGAGT 802

733 CAGTGCCACGCGAAGTCTCGGCTGCCCCAAGAGGCCAGGGCTGGCGCTGCCCTCTGAGCC 792

803 GGAGAGGGACCCACACAGGCAGGTGTACCAACCCCATCAGGCCAAATCATGGGTGCCAAG 862

793 GGAGGGAGCGCCGTTGGGACAGGGTCTTGGGCCCAACCCGGGACGAGCGGTGGACCGAG 852

863 TCCTGCTGGTTCGCCGAGGTGGCTACT-----GCAGAGAAAGTTGTGTTCTAA 913

853 TGACCGTGGTTCTCTGTGTGTGTCACTGCCAGACCCGCCGAAGAGCCACTCTTTGGA 912

914 AGGAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGTGTGCTGTATAACACAGCCCGAG 970

913 GGGTGGCTCTCTGGCAGCGGCACATCCACCCATCGTGGGCGCCAGCACACACGCGG 972

971 CTCACAT---CTCTGCTGCACCAACCCCGCAAAATGCTTTGAGTCTAGGCCCATTTAT 1027

973 CCCCCCATCCACATCGCGGCCACCACTCCCTGGGACACGCTGTCTCCCGGTGTACGC 1032

1028 TGAGACCCAGACATTTCTTTACTCTCAGGGAGATGGCCAGAGCGTCTTAACCCCTCAT 1087

1033 CAGAGCAACAGCACTTCTCTACTCTCTCAGGCGA---CAAGAGCAGCTCGCGCCCTCT 1089

1088 CTTACTCAGCAACTCTCAGAGCTTAATTGACTGGGGCCAGGACCTGGTGGAGATCATCTT 1147

1090 CTTACTCAGCTCTCTGAGGCCCGAGCTGACTGGCGCTCGAGGGCTCGTGGAGACCATCTT 1149

1148 TTGGGCTCAAGGCTTAGACATCAGAACATCTTGAGGACACACCGTCTATCGCGTCG 1207

1150 TCTGGTTCCAGGCCCTGGATGCCAGGGAATCCCGCAGGTTGCCCCGCCCTGCCCGACG 1209

1208 ATACTGGCAGATGGGCCCTGTTTCCAAACAGCTGTCTGTGTGAACCATGACAGATGCCAATA 1257

1210 CTACTGGCAATATGGGCCCTTTGTTCTGAGAGTGTCTTGGGAACCAACGCGCATGCCCTA 1269

1268 TGTTCAGACTCTTCAGGTCAATTCACAGTTTGGAAACAGCAAAACCAACAGGTGACAGATGC 1327

1270 CGGGGTGCTCTCTAAGAGGCACCTCCCGCTGCGAGTGGGGTCACCCCAGCAGCGCGGT 1329

1328 CT-----TGAAACACAGGCC 1342

1330 CTGTGCCCGGAGAGCCCGCAGGCTCTGTGGCGGCCCGGAGGAGGACACAGACCC 1389

1343 ACCGCACTCATGGATTGCTCCGCTGCAACAGAGTCCCTGGCAGGTATATGGTTTTCT 1402

1390 CCGTGGCTTGTGACAGTCTCCGCCAGCAGACAGACCCCTGGCAGGTGTACGGCTTCGT 1449  
1403 TCGGGCTCTCTCTGCAAGGTGGTGTCTGTAGTCTCTGGGGTACCAAGCACAATGAGCG 1462  
1450 GGGGGCTCGCTCGCGCGCTGTGCCCCAGGCCCTCTGGGGCTCCAGGCACAACGAAAG 1509  
1463 CGCTCTCTTTAAGAACTTAAAGAACTTATCTCTGTTGGGAAATACGCAAGCTATCACT 1522  
1510 CGCTTCTTCAGGACACCAAGAACTTATCTCTCTGGGAAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGGAACCTGATGTGGAAGATGAAGTGAAGATGACCTGGTTCGCGACAGCCCGGG 1582  
1570 GCAGGAGCTGAGGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGAGCCGAGG 1629  
1583 GAAGGACCTGTCCCGCTGCAGACACCGTCTGAGGAGAGGATCCCTGGCTACGTTCTCT 1642  
1630 GGTGGCTGTGTTCCGGCGCAGAGACCGCTCTGCTGAGGAGATCTTGGCCAAGTTCTCT 1689  
1643 GTTCTGGCTGATGGACACATACGTGGTACAGCTGCTTAGGTCATTTCTTTATACACAGA 1702  
1690 GCATCGCTGATGAGTGTGTAGTGTGCTGAGCTGCTCAGGTCCTTTTATGTACCGGA 1749  
1703 GAGCATTTCAGGAAGACAGGCTCTTCTTACCGTAAGAGTGTGGAGCAAGCTGCA 1762  
1750 GACCACTTTCAGGAAGACAGGCTCTTCTTACCGGAAGAGTGTGTGGAAAGTTGCA 1809  
1763 GAGCATTGGAGTCAGGCAACACCTTGAGAGAGTGGGCTACGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTGNAATCAGACAGCATTTGAAGAGGTGACGCTGGGAGAGCTGTGGAGCAGA 1869  
1823 GGTGAGGATCACCAGGACACTGTGGTACGCAATGCCATCTGCGAGACTGGCGCTTCATCCC 1882  
1870 GGTGAGGACCATTCGGGAAGCCAGGCCCGCTGTGACGTCGACACTCCGCTTCATCCC 1929  
1883 CAAGCCCAAGCGCTCGCGCCCATTTGTGAACATGATTATAGCATGGGTACACAGAGTTT 1942  
1930 CAAGCTGACGGCTCGCGCCCATTTGTGAACATGACTACGTGCTGGAGACCAAGATT 1989  
1943 GGGCAGAGGAAGAGCCAGCCAGATTTCAACCAGCGTCTCAAGACTCTCTTCATGATGCT 2002  
1990 CCGCAGAGAAAGAGGGCGGAGCGCTCTCACTCGAGGGTGAAGCACTGTTCAGCGTGT 2049  
2003 CAACTATGAGCGGCAAAATCATCTCACTTATAGGGTCTCTGTACTGGGTATGATGA 2062  
2050 CAATGAGAGCGGCGCGGCCCGCGCTCTCGGGCGCTCTGTGCTGGGCCCTTGGACGA 2109  
2063 CATCTACAGGACCTGGCGGGCTTTGTGCTGCTGTGGTGTCTGACACCAAGACACCCAG 2122  
2110 TATCCACAGGGCTGGCGCACCTTCGTGCTGCTGTGCGGGCCAGGACCGCGCCTGA 2169  
2123 GATGTACTTTGTAAGCAGATGTGACGGGGCTTATGATGCCATCCCGCAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGGTGTGACGGGGCGTACACACCATCCCGCCAGGACAGGCT 2229  
2183 GGTGAGGTTGTTCGCAATATGATCAGGCATCTCGGAGAGCAGTACTGTATTCGCCAGTA 2242  
2230 CACGAGGTCATCGCCAGCATATCA--AACCACAGAACAGTACTGCTGCTCGGTA 2286  
2243 TGCAGTGGTCCGAGAGATAGCCAGGCCAAGTCCAAGTCTCTTTAGGAGACAGGTCAC 2302  
2287 TGCGGTGGTCCAGAGGCGCGCCATGGGCAGCTCGCGAAGGCTTCAAGAGCCACGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCTTAAAGCATCTGCAGGATTCCGA 2362  
2347 TACTTTGACAGACCTCCAGCCGTAATGCGACAGTGTGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGCATGAGGAATCCGTTGTGATATGAGCAGAGCATCTTATGATATGAGACAG 2422  
2401 GACCAAGCCGCTGAGGATGCGCTGTGATCAGCAGAGAGCTCTCCCTGATATGAGGCCAG 2460  
2423 CACGAGCTGTTGACTCTTCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTCGACGCTCTTCTAGCTTATGTGCCACCAAGCCGCGGCGATCAGGGG 2520

QY 2483 CAGGTGCTATACGAGTGCAGGGATCCCGCAGGGCTCCAGGCTATCCACCGCTGCTCTG 2542  
DB 2521 CAAGTCTTACGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
QY 2543 CAGTCTGTGTTTCGGAGACATGAGGAAACAAGCTGTTTGTCTGAGGTGCAGCGGATGGGTT 2602  
DB 2581 CAGCTGTGCTACGGGACATGAGGAAACAAGCTGTTTCCGGGATTCGGCGGACGGCT 2640  
QY 2603 GCTTTTACGTTTGTGATGATCTTCTGTTGGTGAAGCTCCTACCTTGAACCAAGCAAAAC 2662  
DB 2641 GCTCTGCGTTTGGTGAATGATTTCTTGTGGTGACACCTCACCTCACCCACGCAAAAC 2700  
QY 2663 CTTCTCAGCACCTGTGTCATGCGGTTCTGAGTATGGTGTGATATAAATTTGAGAA 2722  
DB 2701 CTTCTCAGGACCTGTGTCGAGGTGCTTCTGATATGGTGTGGTGAATTCGGAA 2760  
QY 2723 GACAGTGTGAATTCCTCTGTGAGCCTGTGAACCTGGGTGGTGCAGCTCCATPACAGCT 2782  
DB 2761 GACAGTGTGAATTCCTCTGTGAGGACGAGGCGCTGGTGGCACGGCTTTTGTTCAGAT 2820  
QY 2783 GCCTGCTCACTGCGCTGTTTCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
DB 2821 GCGGCGCCACGCGCTATTTCCCTGCTGGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACCTTCCAGAG 2902  
DB 2881 GCAGAGCGACTACTCCAGTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTAAAGCTGGGAAGACCATGCGGAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2962  
DB 2941 CGGCTTCAAGCTGGGAGGACATGCGCTGCGCAACTCTTTGGGGTCTTGGGGTGAAGTG 3000  
QY 2963 TCACGCTCTATTCTTAGACTTGCAGGTGAACAGCCTCCAGACAGCTCTGCATCAATATATA 3022  
DB 3001 TCACAGCCTGTTCTGATTTGACAGTGAACAGCTCCAGCGTGTGCAACCAATCTA 3060  
QY 3023 CAAATCTTCTGCTTACGGCTACAGGTTCCATGATGTGATGTGATTCAGCTTCCCTTTGA 3082  
DB 3061 CAAATCTTCTGCTTACGGCTACAGGTTTCAGCATGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
QY 3083 CCAGCGTGTGAGGAAGAACCTCACTTCTTCTGGGCTCATCTCCAGCCCAAGCATCCTG 3142  
DB 3121 TCAGCAAGTTTGAAGAAACCCCACTTTTCTGCGGCTCATCTCTGACACGCGCTCCCT 3180  
QY 3143 CTGCTATCTATCTGAGGTCAAGATCCAGGAATGACACTAAAG-----GCCTC 3193  
DB 3181 CTGCTACTCTCATCTGAAAGCAAGAACGAGGATGTGCTGGGGGCCCAAGGGCGCGC 3240  
QY 3194 TGGCTCTTCTCTGAGCGCGCACATTTGGCTCTGCTACAGGCTTCTGCTGCTCAAGCT 3253  
DB 3241 CGGCGCTCTGCGCTCCGAGGCGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
QY 3254 GGTGCTCATCTGCTGATCTCAAAATGTCTCTGGGACCTCTCAGGACAGCCCAAAACT 3313  
DB 3301 GACTCGACACCGTGTCACTAGCTGCGCACTCTCTGGGTCACTCAGGACAGCCAGACGCA 3360  
QY 3314 GCTGTGCGGAGTCTCCAGAGGCGCAATGACCATCTTAAAGCTGCACTGACCCAGC 3373  
DB 3361 GCTGAGTGGAGTCTCCGGGACGAGCTGACTGCTGAGGCGCGCAGCCAAACCCCGC 3420  
QY 3374 CTTAGCAGCAGACTTTCAGACCATTTTGGACTAA 3407  
DB 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

RESULT 4

AAZ20279

ID AAZ20279 standard; cDNA; 4015 BP.

XX AAZ20279;

XX AC

DT 17-JAN-2000 (first entry)

Human telomerase reverse transcriptase (hTERT) cDNA.

Telomerase reverse transcriptase; human; hTERT; cell proliferation; cancer; ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 56..3454  
/\*tag= a

W09950386-A2.

07-OCT-1999.

31-MAR-1999; 99WO-US07097.

31-MAR-1998; 98US-0052864.

03-AUG-1998; 98US-0128354.

(GERO-) GERON CORP.

Morin GB;

WPI; 1999-610842/52.

P-PSDB; AAY32090.

New catalytic polypeptide and polynucleotide, useful for increasing catalytic activity in a cell.

Disclosure; Fig 2; 24pp; English.

This is the nucleotide sequence of cDNA encoding human telomerase reverse transcriptase (hTERT, see AAY32090). Human telomerase is a target for diagnosing and treating diseases relating to cell proliferation and senescence, such as cancer, or for increasing the proliferative capacity of a cell. A claimed method for increasing the proliferative capacity of a vertebrate cell, especially a human or other mammalian cell, involves introducing into the cell a recombinant hTERT polynucleotide encoding an hTERT variant in which residues 192-323, 200-323, 192-271, 200-271, 222-240, 415-450, 192-323 and 415-450, or 192-271 and 415-450 are deleted. A claimed method for reducing telomerase activity in a cell involves introducing a recombinant polynucleotide encoding an hTERT variant having a deletion of amino acids 192-450, 560-565, 637-660, 638-660, 748-764 or 1055-1071. The polynucleotides are obtained by mutagenesis of the hTERT coding sequence.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGGCGCTTGACCAATGACCGCGCTCTCGTGGTCCCGCGGTGGCTCTGCTGG 82  
40 CCGGCGCAACCCCGGATGCGCGCGCTCCCGCTGCGGCGGTGCTGCTGGCG 99  
83 CAGCGCATACCGGAGGTGTGCGCGCTGCGCAACCTTTGTGCGCGCTGGGCGCCGAGGG 142  
100 CAGCGCATACCGGAGGTGTGCGCGCTGCGCAACCTTTGTGCGCGCTGGGCGCCGAGGG 159  
143 CAGCGCGCTGTGCAACCGGCGGACCGCAAGATCAGCGCACTTGTGTTGCCAATGCGCT 202  
160 CTGGCGGCTGTGTCAGCGCGGCGACCCGCGGCTTTCCGCGCGCTGTGTCGCCAGTGCCT 219  
203 AGTGTGCACTGCGGCTCAGCGCTTCCAGCTTCCGCGCGCTTCCGCGCGCTGTC 262  
220 GGTGTGCTGCTGCGGCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
263 ATCCCTGAAGAGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 322

280 CTGGCTGAAGAGCTGTGGTGGCCCGAGTCTGACAGGCTGTGCGAGCGCGCGCGAAGAA 339  
323 CGTGTGGCTTTTGGCTTTGAGCTGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTT 382  
340 CGTGTGGCTTTGCGCTTGGCTTGGACGGGCGCGCGGGGCGCGCGGGGCTT 399  
383 CACTAGTAGCGTGTGCTAGTACTTGGCCCAACACTGTTATTGAGACCCCTGCGTGTCACTGG 442  
400 CACCACCGGTGGCGAGTACTTGGCCCAACACTGTTGACCGACACTGCGGGGAGCGG 459  
443 TGCAATGATGCTACTGTTGAGCGGAGTGGGCGACGACTGCTGCTTACCTGCTGGCACA 502  
460 GCGTGGGGCTGCTGCTGCGCGCGTGGCGAGACGCTGCTGTTTACCTGCTGGCAG 519  
503 CTGTGCTCTTTTATCTTCTGGTGGTGGCCCGCGAGTGTGCTTACAGGTGTGTGGCTCTCCCT 562  
520 CTGCGGCTCTTTTGTGCTGGTGGCTTCCAGCTTGGCTTACAGGTGTGCGGCGCGCT 579  
563 GTACCAAAATTTGTGCCAACACGATATCTGGCCCTCTGTGCTGCTAGTTACAGGCCAC 622  
580 GTACCACTCGCGCTGCTGCTAGTGGCCCGCGCGCGCGCGCGCGCTAGTGGACCCCGAAG 639  
623 CCGACCGTGGCGAGGAATTTCACTAACTTAGTGTCTTCAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAA CGGCGCTGGAACCATAG 672  
683 TCGCCAGGAAGACACCGAAACCCCTGGCTTGCCTCTCGAGGTACAAAGAGGCATCTAG 742  
673 CGTCAAGGAGCGGGTCCCTGCGCTGCGCGCTGCGCGCGGTGCGAGGAGCGCGGG 732  
743 TCTCACCAGTACAGTGTGCTTCAAGTAAAGAGCGCAGATCTATCTGCTCCGAGAGT 802  
733 CAGTGCACGCGGAGTCTGCGTGGCCAAAGAGCGCGCGCTGCGCTGCGCTGAGCC 792  
803 GGAGGAGGAGACCCCAACAGGAGGTGCTACCAACCCCATCAGGCAATCATGGTGCCAAAG 862  
793 GGAGGAGGAGCGCGGTGGCGAGGGTCTGGCGCGCGCGCGCGCGCGCGCGAG 852  
863 TCCTGTGCTGCTGCGCGAGTGGCTTCT-----GCAGGAAGATTTGCTCTTCAA 913  
853 TGACCGTGGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
914 AGGAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGTGCGCTCTGCTGCGAGCGCGCTTCCACCCATCCGTGGCGCGCGCGCGCGCGCGG 972  
971 CTCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCGCGCATCCACATCG 1032  
1028 TGAGACGAGCATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1087  
1033 CGAGACCAAGCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1089  
1088 CCTACTCAGCAACCTCTGAGCTTAACTTGAATGGGCGCGAGGAGTGTGGAGATCATCTT 1147  
1090 CCTACTCAGCTCTCTGAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
1148 TCTGGCTCAGGCGCTGAGCATCAGGACCATCTGCGAGGACACCGCTTATCGCTGCTGCTGCTGCT 1207  
1150 TCTGGCTTCCAGGCGCTGAGTCCAGGAGTCTCCCGCGAGTGTGCGCGCTGCGCGCGCGCGCGCG 1209  
1208 ATCTGGCAGATCGGCGCGCTGTTCCAAAGCTGCTGTTGTAACCATGCGAGGTGCCAATA 1267  
1210 CTACTGGCAATGCGCGCGCTTCTTCTTCTGAGCTGCTTGGAAACCGCGAGTGGCGCTA 1269  
1268 TGTCAAGCTCTCAGTCAATGCGAGTTTCCAAAGCAACCAACAGTGTGACAGATGC 1327  
1270 CCGGCTGCTCTTCAAGACGCACTGCGCGTGGAGTGCCTGCGCTGCGCTGCGCTGCGCTGCTGCT 1329  
1328 CT-----TGAAACACGAGCC 1342  
1330 CTGTGCGCGGAGAGCG 1389

1343 ACCGACCTCATGATTTGCTCCGCTGCAACAGCATGCTCCCTGGCAGGTATATGTTTCT 1402  
1390 CCGTCGCTTGTGAGCTGTCTCGCCAGCAGCAGCAGCCCTCGCAGGTGTACGGTTCGT 1449  
1403 TCGGGCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACCCAGGCAAAATGAGCG 1462  
1450 GCGGCTGTCTGCGCGGTGTGTGCCCCAGGCTCTGGGCTCCAGGCACAGCAAG 1509  
1463 CCGTCTTTAAGAACTTAAAGATTTATCTGTTGGGAAATACGGCAAGCTATCACT 1522  
1510 CCGTCTCTCAGGAACACCAAGATTTATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATGCCACTGGCTCCGAGCAGCCGGG 1582  
1570 GCAGAGCTGACGTGGAAGATGAGCTGCGGACTGCGCTTGGCTGCGAGAGCCCAAG 1629  
1583 GAAGGACCGTGTCCCGCTGCAAGCAGCAGTGTGAGGAGAGGATCTGGTACGTTCT 1642  
1630 GGTGGCTGTGTCCGCGCAGAGCAGCAGTGTGCGTGAAGGATCTCTGTTATGTCA 1689  
1643 GTTCTGCTGATGGACATACCTGCTGAGCTGTAGTCACTTTTATCATCACAGA 1702  
1690 GCATGGCTGATGAGTGTGACGTGCGAGCTGCTCAGTCTTTTATGTCA 1749  
1703 GAGCACAATCCAGAAACAGGCTCTTTCTACCGTAAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAGAAACAGGCTCTTTTCTACCGAAAGAGTGTGTGAGCAAGTTCA 1809  
1763 GAGCATTTGAGTCAAGCAACACCTTGAGAGTGTGCGCTACCGGAGCTGTCAAGAGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGACGCTGCGGAGCTGTGGAAGCAG 1869  
1823 GGTGAGCATCAACAGACACCTTGGGTAGCCATGCCCATCTGACAGCTGCGTTCATCCC 1882  
1870 GGTGAGCAGCATCGGAAGCAGGCGCGCCCTGCTGACGTCCAGACTCGGTTCATCCC 1929  
1883 CAAGCCCAACGCGCTGCGGCCATTTGTGAACATGATTTAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTTGACGGCTGCGGCCGATTTGTGAACATGACTACGTGCTGGAGCCAGAACGTT 1989  
1943 GGGCAGAAAGCAGGCCAGCATTTCAACAGCTCTCAAGCTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCCGAGGCTCTACCTCGAGGTGAAGGACATGTTCAAGCTGCT 2049  
2003 CAATATGAGCGCAAAACATCTCACTTATGCGGTCTTCTGACTGGGTATGAATGA 2062  
2050 CAATAGAGCGGCGCGCGCCCGGCTCTGCGGCTCTGTGCTGGGCTCGACGA 2109  
2063 CATCTAGAGACCTGGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2122  
2110 TATCCACAGGCTTGGCGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
2123 GATGTAATTTTAAAGCAGATGTGACCGGGGCTATGATGCCATCCCCAGGGTAAGCT 2182  
2170 GCTGTAATTTTCAAGTGTGATGTGACGGCGGTACACACCATCCCTCCAGCAGGCT 2229  
2183 GGTGAGGTGTTCGAATATGATCAGGCTACCGGAGCAGTGTGATGCTGCTGCTGCTGCT 2242  
2230 CAGGAGGTATGCGCCAGCATATCA---AACCCAGAACACGTACTGCTGCTGCTGCTGCT 2286  
2243 TGCACTGGTCCGAGAGATAGCAAGGCCAAGTCCCAAGTCTTTTGGAGACAGGTCAAC 2302  
2287 TGCGTGTTCAGAGGCCCGCCATGGGCAAGTCCGCAAGGCTTCAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCAGGCCATATGCGCCAGTTCCTTAAAGCATGTGAGGATTCAGA 2362  
2347 TACCTTGACACAGCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGGCAGGTGCACTGAGGAACCTCGTGTGTCAGAGAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATGCGGTGCTCATGAGCAGAGCTCTCTCTGATGAGGCCAG 2460

QY 2423 CAGCAGCCTGTTTGTACTTCTCTCTGCACTTCTCTGGTCAAGTGTCTGTAAGATTTGGTGA 2482  
Db 2461 CAGTGGCTCTTTTCGACGCTTCTCTACGCTTATGTCGCCACACGCGGTGCGCATCAGGG 2520  
QY 2483 CAGGTCTATACGCAAGTGTCCAGGCAATCCCGAGGGTCCAGGCTATCCAGCTGCTCTG 2542  
Db 2521 CAACTCTACGTCCAGTCCAGGGATCCCGAGGGTCCATCTCTCCAGCTGCTCTG 2580  
QY 2543 CAGTCTGTCTTCCGAGACATGAGAAACAAGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
Db 2581 CAGCTGTCTACGCGCAGCATGGAGAAAGCTGTTTCGGGGATTCGCGGGGACGGGCT 2640  
QY 2603 GCTTTTACGCTTGTGTGATGACTTTCTGTGTGAGCTCAGCTTGGACCAAGCAAAAC 2662  
Db 2641 GCTCTGCTGTTTGTGTGATGATTTCTTGTGTGACACTCACTCACCACCGGAAAC 2700  
QY 2663 CTTCTCAACAACCTTGGTCCATGCGGCTTCTGAGTATGGGTGCAATGATAAATCTGCAGAA 2722  
Db 2701 CTTCTCAAGACCTTGGTCCGAGGTGCTCTGAGTATGGTGGTGGTGAACCTTGGGAA 2760  
QY 2723 GACAGTGTGAACCTTCCCTGCTGGAGCCTGATACCTGGTGGTGGTGCAGCTCCATACCACT 2782  
Db 2761 GACAGTGTGAACCTTCCCTGTAGAAAGCAGGCGCTGGGTGGCAGCGCTTTTGTTCAGAT 2820  
QY 2783 GCTGCTCACTGCTGCTTCTCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2842  
Db 2821 GCGGCGCCACGCGCTATTTCCCTGCTGGGCTGCTGCTGATACCGGACCTCTGGAGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCAGCTTCCAGAG 2902  
Db 2881 GCAGCGGCTACTCCAGCTATGCCGAGACCTCCATCAGACCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAGCTGGGAAAGACCATCGGAAACAAGCTCTGCTGGTCTTGGGTTGAAAGT 2962  
Db 2941 CGGCTTCAAGCTGGGAAAGACCATCGGAAACAAGCTCTTGGGCTTGGGCTGAAAGT 3000  
QY 2963 TACGCTGTATTTCTAGACTTGCAGGTGAAAGCTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGGATTTGCAAGTGAACAGCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAAGATCTTCTGCTTCCAGGCTACAGGTTCCATGATGCTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAGATCTTCTGCTGAGGCTACAGGTTTCAAGCTGCTGCTGAGCTTCCATTTCA 3120  
QY 3083 CAGCGTGTGAGAAAGCTCACTTCTTCTGGGCAATCTTCCAGCCAGGATCTGCTCAAGCT 3142  
Db 3121 TCAGCAAGTTTGAAGAAACCCCACTTTTCTGCGCTCATCTCTGACAGGCTCCCT 3180  
QY 3143 CTGCTATGCTATCTTGAAGTCAAGATCCAGGATGACACTAAG-----GCTC 3193  
Db 3181 CTGCTACTCCATCTGAAAGCCAGGACGAGGATGCTGCTGGGGCCAGGGCGCGC 3240  
QY 3194 TGGCTCTTCTCTCAGAGCGCAGCATTTGGCTCTGCTTACAGGCTTCTGCTCAAGCT 3253  
Db 3241 CCGGCTCTGCTCCGAGGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
QY 3254 GGTGCTGCTATCTGCTGCTTACAAATGTCTCTGGGCTCTGAGGACAGCCCAAACT 3313  
Db 3301 GACTGACACCGGTGCTCACTAGTCCACTCTGGGCTCACTCAGGACAGCCAGCGCA 3360  
QY 3314 GCTGTCCGGAAGCTCCAGAGGCGCAATGACATCTTAAAGCTGAGCTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTCGGAAGCTCCCGGGACGACGCTGACTGCCCTGAGGCGCGAGCCACCCGCG 3420  
QY 3374 CTTAAGCAGACTTTTCAGACCATTTTTCAGCTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGAGCTGA 3454

RESULT 5  
AAZ00724  
ID - AAZ00724 standard; DNA; 4015 BP.  
XX

AAZ00724;

06-OCT-1999 (first entry)

Human telomerase catalytic domain DNA.

Telomerase; catalytic domain; human; quantitation; tumour cell; melanoma; body fluid; metastases; T-cell lymphoblastoma; chronic myeloid leukemia; acute lymphatic leukemia; melanoma; pulmonary carcinoma; colon cancer; breast cancer; ss.

Homo sapiens.

DE19804372-A1.

05-AUG-1999.

04-FEB-1998; 98DE-1004372.

04-FEB-1998; 98DE-1004372.

(DAHM/) DAHM M W.

Dahm MW;

WPI; 1999-431408/37.

Quantifying tumor cells by amplifying mRNA encoding the catalytic subunit of telomerase

Example; Fig 1A-B; 26pp; German.

This invention describes a novel method for the quantitation of tumour cells in a body fluid which comprises (1) enrichment or isolation of tumour cells in the sample, (2) amplification of mRNA from these cells that encodes the catalytic subunit of telomerase and (3) quantifying the amount of amplified mRNA. The method is applied to tumour cells derived from (micro)metastases, e.g. associated with a wide range of tumours such as T-cell lymphoblastoma, chronic myeloid or acute lymphatic leukemia, melanoma, pulmonary carcinoma, cancer of colon or breast etc. This sequence encodes a human telomerase protein catalytic domain.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGCGCTTGACACATGACCCCGCTCTCGTTGCCCGCGGTGCGTCTCTGTCGG 82

40 CCCGCGCACCCCGCGATGCGCGCGCTCCCGCGTCCGAGCGGTGCGTCTGTCGG 99

83 CAGCGGATACCGGGAGGTGTGGCGCTGGCAACCTTTGTGCGCGCTGGGGCCCGAGGG 142

100 CAGCCACTACCGGAGGTGTGCGCTGGCCACGTTCGTGCGCGCTGGGGCCCGAGGG 159

143 CAGCGCGCTTGTGMAACCGGGGACCGAAGATCTACCGACATTTGGTTGCCCAATGCTT 202

160 CTGGCGGCTGTGTGACGCGGGGACCGGGCGGCTTTCCGCGCGTGTGGCCCGAGTGCCT 219

203 AGTGTGCATGCACTGGGGCTCAGACGCTCCACCTGCGGACCTTTCTTCCACAGGTGTC 262

220 GGTGTGCTGCTTGGGACGACGCGCGCGCCCGCGCCCTCTCTTCGCGCAGGTGTC 279

263 ATCCCTGAAGAGCTGTGGCCAGGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322

280 CTGCTGAAGAGCTGTGGCCCGAGTGTCTCAGAGGTGTGCGAGCGCGCGCGAAGAA 339

323 CGTGTGCTTTTGGCTTTGAGCTGCTTAACGAGCGCAGAGGGCGGCTCCCATGGCTTT 382

340 CGTGTGGCGCTTCGCGCTTCGCGCTGCTGAGCTCTGTGGGGTACCGAGCGCAATGAGCG 399

383 CACTAGTAGCGTGTAGTACTTGGCCCAACACTGTTATTGAGACCTGCGTCTCAGTGG 442  
 400 CACCACAGCGTGGCAGCTACTGTCCTCCCAACACGCTGACCGACGACTGCGGGGAGCGG 459  
 443 TGCATGATGCTACTGTTGAGCCGAGTGGGCGACACCTGCTGGTCTACCTGCTGGCACA 502  
 460 GGGTGGGGGCTGCTGTCGCGCGGTGGGCGACGACGTGCTGGTTACCTGCTGGCAGC 519  
 503 CTGTGCTCTTTATCTTCTGGTGCCTCCCGCAGCTGTGCTTACAGGTGTGTGGTCTCCCT 562  
 520 CTGCGGCTCTTTGTGCTGGTCTCCAGCTGTGCGCTTACCAGGTGTGCGGGCGCGCT 579  
 563 GTACCAATTTGTGCCACCAAGATATCTGGCCCTCTGTGTCGCTAGTTACAGGCCAC 622  
 580 GTACCAAGCTGGCGCTGCCACTCAGGCCCGGCCCGCCACACGCTAGTGGACCCGAG 639  
 623 CCGACCCGTGGGAGGAATTTCACTAACTTAGTTCCTTAACAAGATCAAGAGCAGTAG 682  
 640 GCGTCTGGGATGC-----GAAACGGGCTTGGAAACCATAG 672  
 683 TCGCCAGGAGCACCGAAACCCCTGGCCTTGCATCTCGAGGTACAAAGAGGCATCTCAG 742  
 673 CGTCAGGAGGCGGGGTCCTCCCTGGGCTGCGACCCCGGGTGGAGAGGCGGGGG 732  
 743 TCTCACCAGTCAAGTGTGCTTCAAGTAAGAGCCAGATGCTATCTCTGCCGAGAGT 802  
 733 CAGTGCACCGGAAGTCTGCCGTTCGCCAAGAGGCCAGGCGTGGCGCTGCCCTGAGCC 792  
 803 GGAGGAGGACCCACAGCGAGGTGTACCAACCCCATCAGGCAATCATGGGTGCCAAG 862  
 793 GGAGCGGACCGCGTGTGGCGAGGGTCTTGGGCCACCCCGGGCAGGACGCGTGGACCGAG 852  
 863 TCCTGTCTGCTCCCGCGAGTGCCTACT-----GCAGAGAAAGATTGTTCTTTCTAA 913  
 853 TGACCGTGTGTTCTGTGTGTGTCACTTGCAGACCCCGCGAAGAGCCACTCTTTTGA 912  
 914 AGAAAGGTGTGTGACTGAGTCTCTC---TGCGTGGTGTGTGTTAAACAAGCCAG 970  
 913 GGGTGGCTCTCTGGCACGCGCACTCCCGACCATCTCGTGGCGCGCCAGCACCGCGG 972  
 971 CTCACAT---CTCTGCTGTACACACCCCGCCAAATGCTTTTCAGCTCAGGCCATTTAT 1027  
 973 CCCCCATCCACATCGCGGCAACGCTCTCTGGGACAGCGCTTCTCCCGGTTACGC 1032  
 1028 TGAGACAGACATTTCTTTTACTCCAGGGGAGATGGCCAAAGCGTCTTAAACCTCTATT 1087  
 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCTT 1089  
 1088 CTTACTCAGCAACCTCCAGCTTAACTTGACTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147  
 1090 CTTACTCAGCTCTCTGAGGCCAGCGCTGACTGCGCTCGAGGCTCTGTGGAGACATCTT 1149  
 1148 TCTGGGCTCAAGGCTTAGACATCAGGACCACTCTGCGAGACACACCTCTATCGCGTCG 1207  
 1150 TCTGGGTTCCAGGCCCTGATGCCAGGACTCCCGCAGGTGTCGCCGCTGCCCGCAGCG 1209  
 1208 ATACTGGCAGATGGCGCCCTGTCCACAGCTGTGTGTGACCATGACAGTGCACATA 1267  
 1210 CTACTGCAAAATCGGCCCTGTGTTCTGGAGTGTGTGGAAACACGCGCAGTGTGCCCTA 1269  
 1268 TGTACAGCTCTCAGGTACATTTGCAAGTTCGAAACAGCAACCAACAGGTGACAGATGC 1327  
 1270 CGGGTGTCTCTCAAGACGACATGCCCCGCTGCGAGTGTGCGGTCAACCCAGAGCGGTGT 1329  
 1328 CT-----TGAAACACAGGCC 1342  
 1330 CTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGCCCGCAGGAGGAGGACACAGACCC 1389  
 1343 ACCGACCTCATGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
 1390 CCGTGGCTGTGACGCTGCTCCGACAGCAGCAGGACCCCTGGCAGGTGTACGGCTTGT 1449  
 1403 TCGCGGCTGTCTCTGCAAGGAGGTGTCTGTCTGTCTGTGGGGTACCGAGCGCAATGAGCG 1462

1450 GCGGGCTGCTGCGCGGTGCTGCGCCAGGCCCTCTGGGGCTCCAGGCACAACGACG 1509  
1463 CCGCTTCTTTAAGAACTTAAAGAACTTCACTCTCGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGACACCAAGAACTTCACTCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGGAACATGATGTGGAAATGAAATGAGGATGTCACCTGCTCCGAGCAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAATGAGCGGTGCGGGAAGTCTGCTTGGCTGCGCAGAGCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTCCAGAGCAGCGTCTGAGGAGAGATCCTCGCTPACGTTCT 1642  
1630 GATTGGCTGTGTTCCGCGCAGAGCAGCGTCTGCTGAGGAGATCTGGCCAGTTCT 1689  
1643 GTTCTGCTGATGAGACACATACGTGGTACAGCTGTAGTGTATCTTTTATCATCAGAGA 1702  
1690 GCATGCTGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1749  
1703 GAGCACATTCAGAGAAACAGGCTCTTCTTACCTGTAAGAGTGTGGAGCAAGTCA 1762  
1750 GACCACTTTCAAGAAACAGGCTCTTCTTACCTGTAAGAGTGTGGAGCAAGTCA 1809  
1763 GAGCATTTGAGTCAAGCAACCTTGGAGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGT 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGTTGCAAGTGTGGAGTGTGGAGTGT 1869  
1823 GGTGAGCATCAGCAGACACCTGGTGTAGCCATCTCCAGACTGCGTTCATCCC 1882  
1870 GGTGAGCATCAGCAGACACCTGGTGTAGCCATCTCCAGACTGCGTTCATCCC 1929  
1883 CAAGCCCAAGCGCTGCGGCCATTTGAAACATGAGTATAGCATGGTGTACAGAGCTTT 1942  
1930 CAAGCTTGAAGGCTGCGGCCATTTGAAACATGAGTATAGCATGGTGTACAGAGCTTT 1989  
1943 GGGCAGAGGAGCAGAGCCAGCATTTTCAAGAGCTGTCAAGACTCTCTTCAAGAGCT 2002  
1990 CCGCAGAGAAAGAGGCGCAGCGTCTCACTCGAGGTTGAAGGCACTGTTCAGCGTCT 2049  
2003 CAATATGAGCGGCAAAACATCTCACTCTTATGAGGCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATCAGAGCGCGCGCGCCCGCCCTCTGGGGCTCTGTGCTGGCGTGGACGA 2109  
2063 CATCTACAGACCTGCGCGCTTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2122  
2110 TATCCAGAGGCTTGGCGCACTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2169  
2123 GATCTACTTTTGAAGCAGATGTGACCGGGGCTATGATGCCATCCCGCAGGTAAGCT 2182  
2170 GCTGTACTTTTGAAGTGTGATGTGACGGCGGCTAGCACCATCTCCCGCAGAGCT 2229  
2183 GGTGAGGTTGTTGCCAATATGATCAGGCACTCGGAGACAGTACTGTATCCGCCAGTA 2242  
2230 CACGAGAGTCTATCCAGCATCATCA...AACCCAGAAACAGTACTGCGTGGCTCGGTA 2286  
2243 TGCAAGTGTCCGAGAGATAGCAAGCCAGTTCACAGTCTCTTTAGAGACAGTCTAC 2302  
2287 TGCCGTGTCCAGAGGCGCCATGAGGACGTCGCGAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCAGCCATACATGAGGCGAGTCTCTTAAAGCATCTCAGAGTTTCA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGACAGTTCGTGGCTCACCTCGAG...GA 2400  
2363 TGCCAGTGTCACTGAGGACCTCGCTGTGATCAGCAGACGATCTCTATCAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGAGATCGCGTGTGATCAGCAGAGAGTCTCTCTCTCAATGAGGCCAG 2460  
2423 CAGCAGGCTGTTTCACTCTCTCTGCACTTCTCTGCGCTCAAGTGTCTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTCTGACGCTCTCTCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGTGTCTATCAGGAGTGCAGGCGATCCCGCAGGCTTCAGGCTATCCACCTGCTCTG 2542

Db 2521 CAACTCTACGTCCAGTCCAGGGATCCGAGGGCTCCATCTCTCCAGCGTCTCTG 2580  
Qy 2543 CAGTCTGTGTTTCGAGACATGAGAAACAAGCTGTTTCTGAGGTGAGCGGGATGSGTT 2602  
Db 2581 CAGCTGTGCTACGGCGCATGAGAAACAAGCTGTTTTCGGGGGATTCGGCGGACGGGCT 2640  
Qy 2603 GCTTTTACGTTTGTGTGATGATCTTCTGTTGGTGAAGCTCCTTGGACCAAGCAAAAC 2662  
Db 2641 GCTCTGCGCTTTGGTGGATGATTTCTTGTGTGACACTCACCTCACCGACCGGAAAC 2700  
Qy 2663 CTTCTTCAGACACCTGCTTCATGCGGTTTCTGAGTATGAGTATGAGTATGAGTATGAGT 2722  
Db 2701 CTTCTTCAGACACCTGCTTCATGCGGTTTCTGAGTATGAGTATGAGTATGAGTATGAGT 2760  
Qy 2723 GACAGTGTGAATCTTCCCTGTGAGCCTGTTACCTGGTGGTGCAGCTCCATACCAAGCT 2782  
Db 2761 GACAGTGTGAATCTTCCCTGTGAGACAGGCGCTTGGTGGCAGGCTTTTGTTCAGAT 2820  
Qy 2783 GCTGTCTCACTGCTGTTTCCCTGTGAGTGTGGTGTGCTGTGACACTCAGACTTTTGGAGT 2842  
Db 2821 GCGGCGCCAGGCTTATCCCTGTGCGGCTGCTGTGATACCGGACCTTGGAGT 2880  
Qy 2843 GTTCTGTGACTACTCAGGTTATGCGGACACTCAATTAAGACGAGCTCACCTTCCAGAG 2902  
Db 2881 GCAGAGCGACTACTCCAGCTATGCCGAGCTTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
Qy 2903 TGTCTTCAAGCTGGGAGACCATCGGAAACAAGCTCCTGCTGCGTCTTGGGTTGAAAGT 2962  
Db 2941 CGGCTTCAAGCTGGGAGACCATCGGCTGCGAACTCTTGGGCTTGTGGCTGAAGT 3000  
Qy 2963 TCAAGCTTATTTTCTGAGTTCAGGTTGAGGCTTCCAGACTTCCAGACTTCCATATATA 3022  
Db 3001 TCACAGCTGTTTCTGAGTTCAGGTTGAGGCTTCCAGACTTCCAGACTTCCATATATA 3060  
Qy 3023 CAAAGTCTTCTGCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGAT 3082  
Db 3061 CAAAGTCTTCTGCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGAT 3120  
Qy 3083 CCAGCTGTTTGAAGAAACCTCACATTTCTTGTGGCAATCATCTCCAGCCAGCATCTG 3142  
Db 3121 TCAGCAAGTTTGAAGAAACCTCACATTTCTTGTGGCAATCATCTCCAGCCAGCATCTG 3180  
Qy 3143 CTGCTATGCTATCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGAT 3193  
Db 3181 CTGCTATCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGATGAT 3240  
Qy 3194 TGCTCTCTTCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGATGAT 3253  
Db 3241 CGGCTCTTCTTCAAGCTTCAAGTTCAGTTCATGATGATGATGATGATGATGATGATGAT 3300  
Qy 3254 GGTGCTCATTTCTGCTCATCTCAAAATGCTCTCTGAGGACTCTGAGGACAGCCCAAACT 3313  
Db 3301 GACTGACACCGGTGCTCACCTAGTCCACTCTGAGGCTCACTGAGGACAGCCCAAGCA 3360  
Qy 3314 GCTGTGCGGAGAGCTCCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTCCGAGAGCTCCGAGGACGAGCTGACTGCTGCTTGGAGCCGAGCCAGCCG 3420  
Qy 3374 CTTAAGCAGAGCTTTTCAAGCAATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGAGTGA 3454

## RESULT 6

AAH45901  
ID AAH45901 standard; DNA; 4015 BP.

XX

AAH45901;

XX

DT 06-SEP-2001 (first entry)

XX

DE Human hTERT gene.

XX





503 CTGTGCTTTATCTTCTGTGTCGCCCCCAGCTGTGCTACCAAGTGTGTGGTCTCCCT 562  
520 CTGGCGCTCTTTGTGTGTGTGCTCCAGCTCGCTCTACCAAGTGTGTGGGCGCCCGCT 579  
563 GTACCAAAATTGTGCCACCAAGATATCTGGCCCTCTGTGTCCGTAGTTACAGCCCAAC 622  
580 GTACCAAGTGGGCTGCTCACTCAGGCGCGGCCCGCCACACGCTAGTGGACCCCGAAG 639  
623 CCGACCGTGGGCAAGGAATTTCACTAACCTTAGTTCTTACACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGGCTGGAACCATAG 672  
683 TCGCCAGGAAGACCCGAAACCCCTGCTTCATCTCGAGGTACAAAGAGGCACTCTAG 742  
673 CGTCAGGAGGCGGGGTCCCTCTGGGCTGCCAGCCCGGGTGGAGGAGCGCGGGG 732  
743 TCTCACAGTACAAGTGTGCTTCAGGTAAAGAGGCGAGATGCTATCTGTCCGAGAGT 802  
733 CAGTGCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCCGAGCGTGGCGCTGCCCTTAGGCC 792  
803 GGAGGAGGAGCCCGACAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGTGGCAAG 862  
793 GGAGCGAGCGCGTGTGGGAGGGTCTGGGCCACCCGCGGACGAGCGCTGGACCGAG 852  
863 TCTGTCTGCTGCCGCCCGAGTGGCTACT-----GCAGAGAAAGATTTGTCTTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCACTGCCAGACCGCGGCAAGAGCCACCTCTTTGA 912  
914 AGGAAGGTGTCTACCTGAGTCTCTC-----TGGGTGGTGTGTGTAAACACAGCCAG 970  
913 GGTGTGCTCTCTGGACCGGCACTCTCCACCCATCGTGGGCGGCGACACACAGCGGG 972  
971 CTCACAT-----CTCTGTGTCAACCCCGCCAAAATGCTTTAGCTCAGGCCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCGACACAGTCTCTGGACACGCTTGTCCCGGTGTAGC 1032  
1028 TGAGACAGATTTCTTTACTTCCAGGAGATGGCCAGAGCGTCTAAACCCCTCAT 1087  
1033 CGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCCAGCTAACTTGACTGGGGCGAGAGACTGGTGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCCAGCTGCTGAGCTGAGGCTCGGAGACCATCTT 1149  
1148 TCTGGCTCAAGGCTTAGACATCAGACACATCTGTGAGGACACACCGTCTATCGGTG 1207  
1150 TCTGGTTCCAGGCTCTGATGTCAGGAGTCTCCCGCAGGTTGCCCGCTCGCCAGCG 1209  
1208 ATACTGGCAGATCGGCCCTGTTTCCAAAGCTGTGGTGAACCATGACAGTGGCAATA 1267  
1210 CTACTGGCAATGGGCCCTGTTTCTGGAGCTGCTTGGGAACACAGCGCATGCCCCA 1269  
1268 TGTCAAGTCTCTCAGGTCACTTGCAGGTTTGAACAGCAAAACCAAGGTGACAGTGC 1327  
1270 GGGGTGCTCTCTAAGACGCACTGCCCGCTGCGAGTGTGCGTCAACCCAGCAGCGGT 1329  
1328 CT-----TGAAACACAGCC 1342  
1330 CTGTGCCGAGAGAGCCCGAGGCTCTGTGGCGGCCCGAGAGGAGGACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CGTCCGCTGTGTGAGTGTCTCCGCGAGCAGCAGCGCTCTGGCAGGTGTAGGCTTGT 1449  
1403 TCGGCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACGAGGACATAGAGG 1462  
1450 GCGGCTCTGCTGCGCGGCTGTGTCCTGAGGCTCTGGGCTTCCAGGACAAAGAAAG 1509  
1463 CGCTTCTTTAAGAACTTAAAGAGTTTCACTCTGTTGGGAAATACGCAAGTATCACT 1522  
1510 CGCTTCTCAGGAACACCAAGAGTTCACTCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1569  
1523 GCAGGAATGATGTGGGAAGATGAAGATAGAGGATTTGCCACTGGCTCGGACAGCCCGG 1582

1570 GCAGGAGCTGAGCTGGAGATGAGCTGCGGAGCTGCGCTTGGCTGGCAGAGCCCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGAGAGACACCGTCTGAGGAGAGGATCTTGGCTAGTTCTCT 1642  
1630 GGTGCTGTGTTCGGCGCGAGAGACACCGTCTGCGTGAAGAGATCTTGGCCAAAGTTCT 1689  
1643 GTTCTGGCTGATGAGACATACGTGTATGAGCTGTCTAGGTCACTTTTATACACAGA 1702  
1690 GCTGCTGATGATGATGTGTAGCTGTGAGCTGTCTAGGTCTTTTATATGTCAGGA 1749  
1703 GAGCACTTCCAGAAAGAACAGGCTCTTTCTAGCGTAAAGATGTGTGAGCAAGCTGCA 1762  
1750 GACCACTTCAAAGAAACAGGCTCTTTTCTACCGAAAGATGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACACTTGTAGAGATGGCTACGGGAGTGTCTCAAGAGGA 1822  
1810 AAGCATTTGAAATCAGACAGCACTTGAAGAGGTGTGAGCTGGGAGCTGTTCGGAAGCA 1869  
1823 GGTCAAGCATCACAGGACACCTGTGTAGCCATGCCATCTGCAGACTTGCCTTTCATCCC 1882  
1870 GGTCAAGCATCGGAGACCGAGCCCGCTGTGCTGAGTCCAGACTCCGCTTTCATCCC 1929  
1883 CAAAGCCAAAGCGCTGCGGCCCATTTGTGAACATGATTTATAGCATGGTATGAGAGCTTT 1942  
1930 CAAAGCTGAGCGCTGCGGCCGATTTGTGAATGAGTACGCTGTGGAGCCAGAGCTT 1989  
1943 GGCAGAAAGAGCAGGCGCCAGCATTTCAACAGCGCTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAAAGAGGCGCGAGCGCTCTCACCTCGAGGGTGAAGCACTGTTTCAGCGTCT 2049  
2003 CAACTATAGCGGACAAACATCTCTCACTTATGGGCTCTTCTGTATCTGGTATGATGA 2082  
2050 CAACTACGAGCGGCGCGCGCCCGCTCTCTGGCGCTCTCTGTCTGGGCTTGGAGCA 2109  
2063 CATCTACAGACTGCGGCGCTTTGTGCTGCTGTGCTGTGCTGTGACACAGACCCAG 2122  
2110 TATCCAGAGGCTTGGCGACCTTGTGCTGGTGTGGGCGCCAGGACCCGCGCTGA 2169  
2123 GATGTACTTTGTAAGGAGATGTGACCGGGCTTATGATGCCATFCCCGCCAGGTAGCT 2182  
2170 GCTGTACTTTGTAAGGTGATGTGACGGGCGCTTACGACACCATCCCGCCAGGACGCT 2229  
2183 GGTGAGGTTGTGTCCATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CAGGAGGTCTACGCGCAGCATCATCA---AACCCAGAACACAGTACTCGGTGCTGCTGA 2286  
2243 TGCAGTGCTCCGAGAGATAGCCAAAGGCCAAAGTCCCAAGTCTTTAGGAGACAGGTCA 2302  
2287 TGCGTGTTCAGAAAGCGCCCATGGGACGCTCCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CAGCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAAGCATCTGCGAGGATTGGA 2382  
2347 TACCTTGACAGACCTCCAGCCGCTACATGCGACAGTTCGTGGCTCACCTGCAAGG 2400  
2363 TGCAAGTGCATGAGAACTCCGTTGTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACAGCCCGCTGAGGATGCGCTGTCTCATCGAGCAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCAGCTCTCTCTAGCTTCTATGTCACCAAGGCTTCCGCTGCTGCTG 2520  
2483 CAGTGTCTACGAGTGCAGGAGTCCCGAGGCTCCAGGCTTATCCAGCTCTCTG 2542  
2521 CAACTCTAGTTCAGTGCAGGAGTCCCGAGGCTTCCAGGCTTCTCTCTGCTGCTGCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAAGCTGTTTGTGAGGTGCAAGGAGTGGGTT 2602  
2581 CAGCTGTGCTACGGGACATGAGAAAGCTGTTTTCGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTTCTGTTGTGAGCTCTCTGCTGAGCCAGCAAGAAAC 2662

```
2641 GCTCCTCGGTTTGGTGGATGATTTCTTGTGGTGACACCTACCTACCCAGCGGAAAC 2700
2663 CTTCTCAGCACCTCGTCCATGCGGTTCTGATGATGGGTGATGATTAACCTTCAGAA 2722
2701 CTTCTCAGGACCTCGTCCGAGGTGCTCGATGATGGTGGTGAACCTTCGGAA 2760
2723 GACAGTGGTGAACCTCCCTCGTGGAGCTGTATCCCTGGGTGGTGGTCCATACCACT 2782
2761 GACAGTGGTGAACCTCCCTGATAGACAGAGGCCCTGGGTGGCAGGCTTTGTCAGAT 2820
2783 GCTGTCTACTGCTGTTTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2842
2821 GCGGCGCACGGGCTATTCCCTGGTGGGCGCTGCTGCTGGATACCGGACCTCGAGGT 2880
2843 GTTCTGTGACTACTCAGGTTATCCCGAGACCTCAATTAAGACGAGCCTCACTTCCAG 2902
2881 GCAGAGGCACTACTCCAGCTATGCCCGGACCTCCATCAGAGCGAGTCTACCTTCAAC 2940
2903 TGTCTTCAAGCTGGGAAGACCATGGGGAACAGCTCCCTGTCGGTCTTCCGGTTGAAGTG 2962
2941 CGGCTTCAAGGCTGGGAGGAACATGGTCCGAACTCTTTGGGCTTTGGGCTGAAGTG 3000
2963 TCAGGCTATTTCTAGACTTGACGTGACGTGACAGCCTCCAGACAGTCTGATCAATATA 3022
3001 TCAGGCTGTTTCTGATTTGAGGTGACGTGACAGCCTCCAGACGCTGTGACCAATCTA 3060
3023 CAAGATCTTCTCTCAGGCTTACAGGTTCCATGATGATGATTCAGCTTCCCTTTGA 3082
3061 CAAGATCTTCTCTCAGGCTTACAGGTTTACGATGATGATGATTCAGCTTCCCTTTCA 3120
3083 CCAGGCTGTTAGGAAGACCTCACTTCTTGGGCACTATCTCCAGCAGCATTCCTG 3142
3121 TCAGCAAGTTTGAAGAACCCTCACTTCTTGGGCGTCACTCTGACAGGCTCCCT 3180
3143 CTGCTATGCTATCTCAGGTTCAAGATCCAGATGACACTAAAG-----GCTC 3193
3181 CTGCTATCTCATCTGAGGCAAGAACCGAGGATGCTGCGGGCCAGGCGCGC 3240
3194 TGGCTCTCTCTCTCAGGCTTACAGGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253
3241 CGGCTCTCTCTCTCAGGCTTACAGGCTTACAGGCTTACAGGCTTACAGGCTTACAGGCT 3300
3254 GGTGCTCTATCTGCTATCTACAAATGCTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313
3301 GACTGACACCTGCTACCTACGTGCTATCTTGGGCTCACTTACAGGACAGCCCAAGCA 3360
3314 GCTGTGCGGAACTCCAGAGCGCAATGACCATCTCTTAAAGTGCAGCTGACCCAGC 3373
3361 GCTGAGTGGAACTCCCGGGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
3374 CTTAAGCACAGCTTTCAGACCACTTTTGGACTAA 3407
3421 ACTGCCCTCAGACTTCAAGACCATCTCTGACTGA 3454
```

ULT 7

46821

AD46821 standard; cDNA; 4015 BP.

AD46821;

27-JAN-2003 (first entry)

Human telomerase reverse transcriptase (TERT) cDNA.

Carbohydrate antigen; alpha(1,3)galactosyltransferase; alpha1,3GT; TERT;  
transgenic; alpha(1,2)fucosyltransferase; alpha1,2FT; human; enzyme;  
telomerase reverse transcriptase; gene; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 56..3454

```
PT FT /*tag= a
XX XX /product= "Human telomerase reverse transcriptase"
PN XX WO200274948-A2.
XX XX 26-SEP-2002.
XX XX 21-MAR-2002; 2002WO-CA00378.
XX XX 21-MAR-2001; 2001US-277811P.
XX XX (GERO-) GERON CORP.
XX XX Denning C, Clark AJ, Schiff JM;
XX XX WPI; 2002-759995/82.
XX XX P-PSDB; AAE29226.
XX XX Mammalian cells, useful for producing animal tissues with carbohydrate
XX XX antigens that are compatible for transplantation into human patients -
XX XX Disclosure; Page 33-34; 71pp; English.
XX XX The invention relates to animal tissues with carbohydrate antigens that
XX XX are compatible for transplantation into human patients. The mammalian
XX XX cell is inactivated homogeneously for expression of alpha(1,3)galactosyl-
XX XX transferase (alpha1,3GT) gene and comprises a transgene for alpha(1,2)-
XX XX fucosyltransferase (alpha1,2Fm). It is useful for producing animal tissue
XX XX with carbohydrate antigens that are compatible for transplantation into
XX XX human patients. The present sequence is human telomerase reverse
XX XX transcriptase (TERT) cDNA used in the invention.
```

SQ Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 24; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

```
QY 23 CCAGCGCTTGAGCAACAATGACCGCGCTCTCTGTTGCCCGCGTCTCTGTGTGG 82
DB 40 CCAGCGCAACCCCGGATGCGCGCTCCCGCTCCCGTCCGAGCGGTGCTCTGTGG 99
QY 83 CAGCGGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGG 142
DB 100 CAGCGCACTACCGGAGGTGTGCGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGG 159
QY 143 CAGCGCGCTTGTCGAACCCCGGAGCGCGGAGATCTACCGCACTTTGTTGCCCAATGCT 202
DB 160 CTGGCGGCTGTGTGACAGCGCGGAGCGCGGCTTTCGCGCGCTGTGTGCGCGCTGT 219
QY 203 AGTGTGATGCACTGGGGCTCAGAGCTCCACCTGCGAGCTTTCTTCCACAGGTGTC 282
DB 220 GGTGTGCGTGGCTTGGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279
QY 263 ATCCCTGAAAGAGCTGGTGGCGAGGCTTGTGACAGACTCTGCGAGCGCAACAGAGAAA 322
DB 280 CTGGCTGAAGAGCTGTGTGGCGCGAGTGTGACAGGCTGTGCGAGCGCGCGCGCGAGAA 339
QY 323 CGTGTGCGCTTTGGCTTTGAGCTGCTTAAAGAGCGCAGAGGCGCGCTCCCATGCGCTT 382
DB 340 CGTGTGCGCTTTGGCTTTGAGCTGCTTAAAGAGCGCAGAGGCGCGCTCCCATGCGCTT 399
QY 383 CACTAGTAGCGTGGCTAGCTACTTTGCCCAACACTGTTATTGAGACCTCGTGTCTAGTGG 442
DB 400 CACCAACAGCTGGCGAGCTACTTGGCGCGAGTGTGCGAGCGCGCGCGCGCGAGCG 459
QY 443 TCATGTGATGCTACTGTTGAGCGAGTGGCGAGCACTGCTGTCTACTGTGTGGCACA 502
DB 460 GCGGTGGGCGCTGCTGCTGCGCGCGTGGCGAGCACTGCTGTGTGTGTGTGTGTGTGT 519
QY 503 CTGTGCTCTTTATCTTCTGTGGCGCGCGCGCTGTGCTACAGGTGTGTGTGTGTGTGTGT 562
DB 520 CTGCGCGCTCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 579
```

563 GTACCAAAATTTGTGCACACACGGAATCTGTGGCCCTCTGTGTCGGCTAGTTACAGGCCAC 622  
580 GTACCAAGCTCGGGCTTCCCACTCAGCCCGGGCCCGCCACACGCTAGTGGACCCCGAAG 639  
623 CCGACCCGTGGGCAGGAATTTCACTAACCTTAGTTCTTTACAAACAGATCAAGACAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGGCTGGAACCATAG 672  
683 TCGCCAGGAAGACCCGNAACCCCTGGCTTGGCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 CGTCAGGGAGCGGGGCTCCCTCGGCTTGCAGCCCGGGTGGAGAGCGCGGGGG 732  
743 TCTCACCAGTACAAAGTGTGCTTCACTAAGAGGCCAGATCTATCTGTCCGAGAGT 802  
733 CAGTGCACGCCAAGTCTGCTGTCGCAAGAGGCCAGCGTGGCGTGCCTCTGAGCC 792  
803 GGAGGAGGACCCACAGCAGGTGTGTACAAACCCCATCAGCAAAATCATGGTGCCAAG 862  
793 GGAGCGGACCGCGTTGGGCAGGGTCTTGGGCCACACCGGGCAGGACGCGTGGACCGAG 852  
863 TCCTGCTCGGTCCCGAGGTGCTACT-----GCAGAGAAGATTGTCTCTAA 913  
853 TGACCGTGGTTCTGTGTGTGTACCTGCGCAGACCCGCGAAGAGCCACCTCTTTTGA 912  
914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTGCGTGTGTCTGTAAACAAAGCCCGAG 970  
913 GGGTGGCTCTCTGGCAGCGGCCACTCCACCCCATCGTGGGCCGCGCAGCACACGCGGG 972  
971 CTCACAT----CTCTGCTGTACACCGCCCGCCAAATGCTTTAGCTCAGGCGCATTTAT 1027  
973 CCCCCCATCACAATCGCGGCCACCACTGCTCTGGACACGCGCTTGTCCCGGGTATCGC 1032  
1028 TGACACAGACATTTCTTTACTTCCAGGGGAGATGCCAAGAGCGTCTAAACCCCTCATTT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCAGCTTAACCTGACTGGGCCAGGAGACTGGTGGAGATCATTT 1147  
1090 CTTACTCAGCTCTCTAGGCGCCAGCTGACTGCGCTCGAGGCTCGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGACCACTCTCAGGACACACCGTCTATCGCGTCG 1207  
1150 TCTGGGTTTCAGGCCCTCGATGCGCAGGACTCCCGCAGGTTGCGCGCTCGCCCGCAGG 1209  
1208 ATACTGCAGATCGGCCCTCTGTTCACAGCTGCTGTGTAAACCAATGCAGAGTGCATA 1267  
1210 CTACTGGCAAAATCGGGCCCTGTTTCTGGAGCTGCTTGGGAACCAAGCGAGTGGCCCTA 1269  
1268 TGTCAAGTCTCAGGTCAATTCAGAGTTTCGAACAGCAAAACCAAGGTGACAGATGC 1327  
1270 CGGGGTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCACCCCAGCGCGGTGT 1329  
1328 CT-----TGAAACACCAAGCCC 1342  
1330 CTGTGCGCGGGAGNAGCCCGAGGCTCTGTGGCGGCCCGGAGGAGGAGACACAGACC 1389  
1343 ACCGCACTCATGATTTGCTCGCTGACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402  
1390 CCGTGCCTGTGTGAGCTGTCTCGCCAGCAGCAGCAGCCCTGCGAGGTGTACGGTTCGT 1449  
1403 TCGGGCTGTCTGCAAGTGTGTGTCTAGTCTCTGGGTGACAGGCAAAATGAGCG 1462  
1450 GCGGGCTGTCTGCGCGGTGTGTGCCCCAGGCTCTTGGGGCTCCAGGCAACAGACG 1509  
1463 CCGTTCTTTAAGACTTAAGAGTTCTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGTCTCTCAGGAAACCAAGAGTTCTCTCCCTGGGGAAGCATGCAAGCTCTCGCT 1569  
1523 GCAGGAACATGTGGAAGATGAAGTAGGAGATTGCACTGCTCGCTCCGAGCAGCGCGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGTTCGCTGCGCAGGAGCCAGG 1629

1583 GAAGGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGGAGTCTGCTAGCTTCCT 1642  
1630 GGTGGCTGTGTTCGGCCGCGAGAGCACCGTCTCGTAGAGATCTTGGCCAAAGTTCCT 1689  
1643 GTTCTGGCTGTATGGACACATACAGTGGTACAGCTGTTAGTTCATCTTTTACATCAGAGA 1702  
1690 GCACTGGCTGTATGAGTGTGTAGTGTCTGAGCTGCTCAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCATTTCCAGAGAACAGGCTCTTCTTCTACCGTAAGAGTGTGTGGACCAAGCTGCA 1762  
1750 GACCAAGTTTCAAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTTCAGCAACACCTTGAGAGAGTGGCTACGGAGCTGTCAACAAGAGA 1822  
1810 AAGCATTTGAATTCAGACAGCACTTGAAGAGGTTGCACTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTAGGCAATCCAGGACACCTGTGCTAGCCATGCCCATCTGACAGACTGCGCTTCATCCC 1882  
1870 GGTAGGCAAGCATCGGGAAGCAGGCCCGCTGTGACGTCCAGACTTCGCTTCATCCC 1929  
1883 CAAGCCCAACGGCTTGGGCCCATTTGTGAACATGAGTTATAGCATGGTACCAAGCTTT 1942  
1930 CAAAGCTGACGGCTTGGGCCGATTGTGAACATGACTACGTCTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAAGAGGCGCCAGCAATTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGTGAAGCACTGTTCAGCGTCT 2049  
2003 CAACTATGAGCGGACAAACATCTTACCTTATGGGTCTTCTGACTGGGTATGAATGA 2062  
2050 CAACTACGAGCGGGCGCGCGCCCGCTCTCTGGCGCTCTGTGCTGGGCTTGGAGCA 2109  
2063 CATCTACAGGACTTGGCGGGCTTTGTGCTGCGTGTGCGTCTCTGGACAGACACCCAG 2122  
2110 TATCCACAGGCTTGGCGCACTTCTGCTGCGTGTGGGCCCGAGACCCGCGCTGA 2169  
2123 GATGTAATTTTAAAGCAGATGTGACCGGGGCTTATGATGCCATCCCGAGGTAAGCT 2182  
2170 GCTGTACTTTTCAAGTGGATGTGACGGGCGGTACGACACCATCCCGCAGGACAGCT 2229  
2183 GGTGAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGAGGTCTATCGCCAGCATCATCA---AAGCCAGAACAGTACTCGTGTGCTCGTA 2286  
2243 TGCAGTGTCTCGGAGAGATAGCCAAAGCCAAAGTCCCTTTTAGGAGACAGGTGAC 2302  
2287 TGGGTGTCTCAGAAAGCGCGCCCATGGCACGCTCGCAAGGCTTCAAGAGCCACGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAAGCATCTGCAGGATTGAGA 2362  
2347 TACCTTGACAGACCTCCAGCCGTATATGCGACAGTTCTGTGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGTGATGAGAACTCGTGTTCATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGGATGCGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCGAG 2450  
2423 CAGCAGCTGTGTGATCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCGACGCTTCTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGGTGCTATATCGCAGTGCAGGGCATCCCCAGGGCTTCCAGCCCTATCCACCTGCTCTG 2542  
2521 CAAAGTCTACGCTCAGTGCAGGGGATCCCGAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTGTGCTTTCGAGACATGGAGAACAGCTGTTGCTGAGGTGAGCGGGATGGGT 2602  
2581 CAGCTGTGCTACGCGCATGGAGAACAGCTGTTGCGGGGATTCGGGGGACGGGCT 2640  
2603 GCTTTTACGCTTTTGTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2662  
2641 GCTCTGCTTGTGCTGATGATTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2663 CTTCTCTCAGACCCCTGGTCCATGGCGTTCCTGAGTATGGGTGCTGATGATAAACTTGCAGAA 2722

```
2701 CTTCTCAGACCCCTGGTCCGAGGTGTCCTGATGATGCTGCGTGTGATCTGCGGAA 2760
2723 GACAGTGTGAATCTCCCTGTGGAGCCCTGGTACCCCTGGTGTGATGCTGATACAGCT 2782
2761 GACAGTGTGAATCTCCCTGTGAAGACGAGGCCCTGGTGTGATGCTGATGCTGAT 2820
2783 GCGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842
2821 GCGGCCCCAGGCTATTCCCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
2843 GTTCTGTACTACTAGGTATGCTCCAGACCTCAATTAGACGAGCTCACTTCCAGAG 2902
2881 GCAGAGCGACTACTCAGCTATGCTCCGAGCTCCATCAGAGCGAGTCTCACTTCAAC 2940
2903 TGCTTCAAGCTGGAGACCATCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2962
2941 CGGCTTCAAGCTGGAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
2963 TCAGGCTTATTCTAGACTTGCAGTGAACAGCTCCAGACAGTCTGATCAATATATA 3022
3001 TCAGAGCTGTTCTGATTTGAGTGAACAGCTCCAGAGCGTGTGCACCAACATCTA 3060
3023 CAAGATCTCTGCTTCAGGCTACAGGTCCATGATGTGATTCAGTTCCTTTGA 3082
3061 CAAGATCTCTGCTTCAGGCTACAGGTTCACGATGTGTGCTGAGCTCCCATTTCA 3120
3083 CCAGGCTGTTAGGAACCTCACAATCTTCTGGGCTATCTCCAGCCAGCATCTG 3142
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCGTCACTCTGACAGGCTCCCT 3180
3143 CTGCTATGCTATCTGAGGCTCAAGATCCAGGAATGACATTAAG-----GCTC 3193
3181 CTGCTATCTCATCTCTGAAAGCAAGAACAGCAGGATGTGCTGGGGGCAAGGGCGCG 3240
3194 TGCTCTCTTCTCTGAGCGGACATTTGCTGCTACAGGCTTCTGCTGCTCAAGCT 3253
3241 CGGCTCTGCTCTCGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
3254 GGTGTCTATCTGCTATCAATAATGCTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313
3301 GACTCGACACCGGTGCTACCTAGTGCCACTCTCTGGGCTCACTCAGGACAGCCAGGCA 3360
3314 GCTGTGCGGAAGCTCCAGAGCGACATGACCTTAAAGCTGAGCTGAGCTGAGCTG 3373
3361 GCTGAGTCGGAAGCTCCGGGAGAGCGCTGACTGCTGAGGCGCGAGCAACCCGGC 3420
3374 CTTAAGCACAGACTTTTCAAGACCTTTTGAATAA 3407
3421 ACTGCCCTCAGACTTCAAGACCATCTGAGCTGA 3454
```

ULT 8  
97534

ABA97534 standard; DNA; 4015 BP.

ABA97534;

05-APR-2002 (first entry)

Cancer cell discrimination method related human DNA.

Human; telomerase; enzyme; cancer cell discrimination; gene;  
reverse transcriptase; ds.

Homo sapiens.

JP2001309791-A.

06-NOV-2001.

02-MAY-2000; 2000JP-0138250.

```
02-MAY-2000; 2000JP-0138250.
(KANE/) KANEUCHI H.
(KAMI/) KANIMORI M.
WPI; 2002-134853/18.
Discrimination of a cancer cell in a sample tissue, comprises
determining the expression level of a reverse transcriptase component
of telomerase using a hybridization assay -
Claim 2; Page 9-10; 16pp; Japanese.
The present invention relates to a method for the discrimination of a
cancer cell in a sample tissue, which involves determining the expression
level of a reverse transcriptase component of telomerase in a cell
constituting the sample tissue by an in situ hybridization of the mRNA of
the enzyme, and judging a cell showing a higher expression level than
that of the reverse transcriptase component of telomerase in a normal
cell to be a cancer cell. The present sequence is a human DNA used in the
exemplification of the invention.
Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;
Query Match 45.8%; Score 1600.8; DB 24; Length 4015;
Best Local Similarity 68.9%; Pred. No. 0;
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;
QY 23 CCGGCTTTGAGCAATGACCCGCTCTCTGTTGCCCGCGTGGCTCTCTGCTGCG 82
DB 40 CCGGCTTTGAGCAATGACCCGCTCTCTGTTGCCCGCGTGGCTCTCTGCTGCG 99
QY 83 CAGCGATACCGGAGGTGTGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 142
DB 100 CAGCCACTACCGGAGGTGTGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 159
QY 143 CAGGCGCTTTGCAACCCGCGGACCCGAGATCTACCGCACTTTGTTGCCCAATGCT 202
DB 160 CTGGCGCTGTGTGACGCGCGGAGACCCGCGGCTTTCCGCGCTGTGTGCCCATGCT 219
QY 203 AGTGTGATGCACTGGGCTCAGAGCTTCAAGCTTCAAGCTTTCCTTCCACAGGTCT 282
DB 220 GGTGTGCTGCTTGGGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279
QY 263 ATCCCTGAAGAGCTGTGGCGGAGGTGTGCGAGAGCTCTGCGAGCGCAACGAGAGAA 322
DB 280 CTGCTGAAGAGCTGTGGCGGAGGTGTGCGAGAGCTGTGCGAGCGCGCGCGAGAA 339
QY 323 CGTGTGCTTTTGGCTTTGAGCTGCTTAAAGCGGCGAGGCGGCGCTCCCATGCGCTT 382
DB 340 CGTGTGCTTTGCGCTTTCGCGCTGCTGAGCGGCGCGCGCGCGCGCGCGCGCTT 399
QY 383 CACTAGTAGCGTGTGAGCTACTTGGCCNACACTGTTATTGAGACCTGCTGCTGAGTG 442
DB 400 CACCACCGTGTGCGAGCTACTTGGCCNACACTGTTATTGAGACCTGCTGCTGAGTG 459
QY 443 TGCATGATGCTTACTGTTGAGCTGGGCGAGCACTGCTGCTTACCTGCTGCGCACA 502
DB 460 GCGGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 503 CTGTGCTCTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
DB 520 CTGCGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579
QY 563 GTACCAAAATTTGTGCAACAGGATATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCT 622
DB 580 GTACAGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639
QY 623 CCGACCGGTGGGAGGAATTTCACTAATCTTGTGCTTCTTCAACAGATCAAGAGCTAG 682
DB 640 GGTGTGCGATGCT-----GACCGGCTTGAACCATAG 672
QY 683 TCGCCAGGAGCAGCGAAACCCCTGGCTTGGCTTCCATCTCGAGGTACAAAGAGGCTG 742
```

673 CGTCAGGAGCGGGGCTCCCTCGGCTGCGAGCCCGGGTGGAGAGCGCGGGG 732  
743 TCTCACCAGTACAAAGTGTGCTTTCAGCTAAGAGGCGAGATGCTATCTGTCGCGAGAGT 802  
733 CAGTGCACCGAGTCTGCGTGGCCAGAGGCGAGCGTGGCTGCCCTGAGCC 792  
803 GAGAGGAGACCCACAGGAGGTGTATCAACCCCATCAGGCAATCATGGGTGCCAAG 862  
793 GAGAGGAGACCCCGTTGGGAGGGGTCTGGGCCACCCCGGCGAGGAGCGGTGGACGAG 852  
863 TCCTGCTCGGTCCCGAGTGTCTACT-----GCAGAGAAAGATTTGCTTTCAA 913  
853 TGACCGTGGTTTCTGTGGTGTCACTGCCAGACCCGCGAGAGCCACTCTTTGA 912  
914 AGAAAGGTGTGTGACCTGAGTCTCTC-----TGGGTGGGTGTCTGTAAACACAGCCAG 970  
913 GGGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGCGCCAGCACCGCGG 972  
971 CTCACAT---CTCTGTGTACACCCCGCCAAATGCTTTACGTCAGGCGCAATTTAT 1027  
973 CCCCCCATCACATCGGGCCACACAGTCTCTGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTTATCTCCAGGGGAGATGCCAAGAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCATTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCCT 1089  
1088 CCTACTCAGCAACCTCCAGCCTAACTTGACTGGGCGCAGGACGCTGGTGGAGATCATCT 1147  
1090 CCTACTCAGCTCTCTGAGGCCCGAGCTGAGTGGCGTCGGAGGCTCGTGGAGACCATCT 1149  
1148 TCTGGGCTCAAGGCTAGGACATCAGGACCATCTCTGAGGACACACCCGTCTATCGCGTCG 1207  
1150 TCTGGTTCAGGCCCTGGATGCCAGGGCTCCCGCAGAGTTCGCCCGCTGCCCCAGCG 1209  
1208 ATACTGCAGATGGGCCCTGTCTCCAGCTGCTGTGTGACCTGTGACCTAGAGTGCCANTA 1267  
1210 CTACTGGCAATGCGGCCCTGTCTTCTGAGCTGTCTTGGGAAACCAAGCGAGTGGCCCTA 1269  
1268 TGTGAGACTCTCAGGTCTCATTTGAGGTTTCAACAGCAAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGAGCAGCTGCCCGCTGCGAGCTGCGGTCACCCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGGCC 1342  
1330 CTGTGCCCGGAGAACCCAGGCTCTGTGGGCGCCCGCGAGGAGGACACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CCGTGGCTGTGAGCTGTCTGCCAGCAGCAGAGCCCTGGCAGGTGTACGGTTGCT 1449  
1403 TCGGGCTGTCTGTCAAGGTGTGTCTGTAGTCTCTGGGTACAGGCACAATGAGCG 1462  
1450 CGGGCTGTCTGCGCGGCTGTGTGCCCCCAGGCTCTGGGGTCCAGGCAACAAGAACG 1509  
1463 CCGCTCTTTAAGAACTTAAGAGTTCATCTGTTGGGAATACGGAAGTCACT 1522  
1510 CCGTCTCTCAGGAACACAGAAAGTTTCTCTCCCTGGGAGACATGCCAGCTCTGCT 1569  
1523 GCAGGAATGTGTGGAAGATGAAGTAGAGGATGTCACTGGCTCCGAGAGCGCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGGACTGCGTTGGCTGCGGAGAGCCAGG 1629  
1593 GAAGGACCGTGTCCCGCTGAGAGCAGCTCTGTGAGGAGAGGATCCTGGCTACGTTCT 1642  
1630 GGTGGCTGTCTCCCGCCGAGAGCACCGTCTGCGTGGAGGATCCTGGCAGGTTCT 1689  
1643 GTCTGGCTGATGGACACATACGTGGGTACAGCTGCTTAGGTCACTTTTATCATACAGA 1702  
1690 GCATGGCTGATGAGTGTGACGTCTGCGAGCTGCTCAGTCTTTCTTTATGTACGGA 1749  
1703 GAGCACATTCAGAGAACAGGCTCTTCTTACCGTAAGAGTGTGTGAGCAGAGCTGCA 1762

1750 GACCAGCTTCAAANGAACAGGCTCTTTTCTACCGGAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATGTGAGTCAAGCAACACCTTTGAGAGAGTGGGCTACGGAGAGTGTCAACAGAGGA 1822  
1810 AAGCATTTGAATCAGACAGCAGCTTTGAAGAGGCTGAGCTGGGAGCTGTCCGAAGCAGA 1869  
1823 GGTCAAGCATACACAGGACACCTGGCTAGCCATCCCATCTGACAGACTGGGCTTCATCCC 1882  
1870 GGTCAAGCAGCATCGGAGAGCAGGCCGCCCTGCTGACGTCCAGACTCGGCTTCATCCC 1929  
1883 CAAGCCAAACGGCTGCGGCCCATTTGAAACATGAGTATATAGCATGGTATACAGAGCTTT 1942  
1930 CAAGCTGACGGCTGCGGCCGATTTGAAACATGACATACCTGCTGGGAGCCAGACGTT 1989  
1943 GGGCAGAGGAGCAGGCCCGCAGCATTTCAACACAGGCTCAAGACTCTTTACAGCATGCT 2002  
1990 CCGAGAGAAAGAGGGCGGAGCTCTCACTCGAGGTGAAGCAGTGTTCAGCGTGTCT 2049  
2003 CAACTATGAGCGGACAAACACATCTCTACCTTTATGGGTCTTTCTGTACTGGGTATGAATGA 2062  
2050 CAACTACGAGCGGCGCGCGCCCGCTCTCTGGGCGCTCTGTGTGGGCTTGGAGCA 2109  
2063 CATCTACAGGACCTCGGGGCTTTGTGCTGGGTGGCTCTGAGCCAGACACCCAG 2122  
2110 TATCCACAGGGCTGGGCACTCTGCTGCTGGTGGGCCCCAGGACCCCGCCCTGA 2169  
2123 GATGTACTTTGTTAAGGAGATGTGACCGGGGCTTATGATGCCATCCCCAGGGTAAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGGATGTGACGGGCGGTAGACACCATCCCCAGGACAGGT 2229  
2183 GGTGAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGGAGTCTATCGCCAGCATCATCA---AAGCCAGAACACGTAAGTCTGCGTGGGTA 2286  
2243 TGCAGTGTCTCGGAGAGATAGCAAGGCCAAGTCCACAAGTCTTTAGGAGACAGGTCAC 2302  
2287 TGGGTGTCTCAGAGGCGGCCCTAGGCACTGCGCAAGGCTTCAAGACCCAGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTCTTAAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCCGTACATGCAAGTCTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCACATGAGGACCTCGTGTGATCAGAGCAGAGCATCTATGAATGAGAGCAG 2422  
2401 GACACGCCGCTGAGGAGTGGCGTGTATGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTGTGATTTCTCTCTGCTACCTTCTCGGTCAAGTGTCTGTAAGATTTGTTGA 2482  
2461 CAGTGGCTCTTTCGACGCTCTCTTACGCTTCTATGTGCCACACGCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGAGTGCAGGCGATCCCGAGGCTCCAGCTATCCAGCTGCTCTG 2542  
2521 CAACTCTACGTCAGTGCAGGGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTCGAGACATGAGAACAAAGCTGTTCGTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGGCGACATGAGAACAGCTGTTTCGGGGATTTCGGCGGACGCGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTGTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTGTGTGTGACACTCACTCACTCACCCACGCGAAAC 2700  
2663 CTTCTCTCAGCACCCTGCTCCATGGGCTTCTCAGTATGGGTGSCATGATAAACTTCAGAA 2722  
2701 CTTCTCTCAGGACCTGCTGTCGAGGTGTCTCAGTATGGCTGCGTGGTGAATTCGCGAA 2760  
2723 GACAGTGGTGAATTCCTGTGAGGCTGGTACCTGGGTGGTGGAGTCCATACAGCT 2782  
2761 GACAGTGGTGAATTCCTGTGTAAGACAGAGGCTTGGTGGACGCTTTTGTTCAGAT 2820  
2783 GCTGTCTCAGCTGCTTTCTCTGCTGGTGTGCTGTGAGACTCAGACTTTTGGAGGT 2842  
2821 GCGGCGCCACGGCTTATTCCTTGGTGGGCGCTGTGTGTGATACCCGACCTGGAGGT 2880

2843 GTTCTGTGACTACTCAGGTTATGCCCAGACCTCAATTAAAGACGAGCCTCACTTTCAGAG 2902  
|||||  
2881 GCAGAGCGACTACTCCAGCTTATCCCGGACCTCCATCAGACGAGCTCTCAGCTTCAACCG 2940  
|||||  
2903 TGTCTTCARAAGCTGGGAAGACCATCGGAACAAGCTCCTGCGGTTCTGGGTTGAAGTG 2962  
|||||  
2941 CGGCTTCACGCTGGGAGGAAACATGCGTCGCAAACTCTTTGGGTCCTGGGCTGAAGTG 3000  
|||||  
2963 TCACGGTCTATTTCTAGACTTCAGGTGAACAGCCTCCAGACAGTGTGCATCAATATATA 3022  
|||||  
3001 TCACAGCCTGTTTCTGATTTGAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTA 3060  
|||||  
3023 CAAGATCTCTCGCTTCAGCCTACAGGTTCCATGCATGTGTATTCAGCTTCCCTTTGA 3082  
|||||  
3061 CAAGATCTCTGCTGGAGCGGTACAGGTTTACGCATGTGTGCTGACGCTCCCATTTCA 3120  
|||||  
3083 CCACGCGTTAGGAAGAACCTCACTTTTCTGGGCATCATCTCCAGCCAAAGCATCTGT 3142  
|||||  
3121 TCAGCAAGTTTGAAGAAACCCACATTTTCTGGCGTCATCTCTGACACGGCTCCCT 3180  
|||||  
3143 CTGTCTATGCTATCCTGAGGTCAAGATCCAGGAATGACATAAG-----GCCTC 3193  
|||||  
3181 CTGTCTATCCATCTGAAAGCCAAAGAACGAGGATGTGCTGGGGGCCAAAGGCGCGCCG 3240  
|||||  
3194 TGGTCTCTTTCTCTGAAGCCGCCAATTTGGGTCTGTCTACAGGCTTCTCTGTCTCAAGCT 3253  
|||||  
3241 CGGCCCTCTGCCCTCGAGGCGCTGCAGTGGCTGTGCCAACCAAGCATTTCTCTGCTCAAGCT 3300  
|||||  
3254 GGCTGCTCATTTCTGTCTATCTAATAATGTCTCTGGGACCTCTGAGGACAGGCCCAAAACT 3313  
|||||  
3301 GACTCGACACCGGTGTCACTTACGTGCCATCTCTGGGTCATCTCAGGACAGGCCGACGCA 3360  
|||||  
3314 GCTGTGCGGAAGCTCCGACAGCGGCAATGACCATCTTTAAAGCTGCAGCTGACCCAGC 3373  
|||||  
3361 GCTGAGTCGGNAGCTCCCGGGGACGAGCTGACTGCCCTGGAGGCGCGAGCCACCCGCG 3420  
|||||  
3374 CCTAAGCACAGACTTTCAGACCAATTTTGGACTAA 3407  
|||||  
3421 ACTCGCTTCAGACTTCAAGACCACTCTCTGGACTGA 3454

JLT 9  
22474  
ABZ22474 standard: cDNA: 4015 BP.

ABZ22474:

25-MAR-2003 (first entry)

Human telomerase reverse transcriptase encoding cDNA SEQ ID NO:1.

Human; telomerase reverse transcriptase; enzyme; hTERT; chromosome 5; vulvectomy; antitumor; epithelial cell migration promoter; wound; epithelialization; skin wound; lesion; burn; surgical incision; ulcer; epithelial cell; keratinocyte; epidermal; mucosal; gene ss.

**Homo sapiens.**

```
Key      Location/Qualifiers
CDS      56..3454
         /*tag= a
         /product= "human te
```

WO200291999-A2.

21-NOV-2002.

09-MAY-2002. 2002WQ-TIS14867

09-MAY-2001: 2001TTS-289903D

(GPO-) GPONY CDD

Jiang X, Chiu C, Harley CB;  
WPI; 2003-120591/11.  
P-PSDB; ABP56676.

Composition for treating wounds and enhancing epithelialization of a skin surface, comprises vector encoding telomerase reverse transcriptase or telomerized epithelial cells on a microparticle or a matrix -

Disclosure; Page 31-32; 68pp; English.

The present invention describes a pharmaceutical composition (I) comprising a vector encoding telomerase reverse transcriptase (TERT) in an excipient or device, or comprises telomerised epithelial cells on a microparticle or a matrix suitable for topical administration or administration to a wound site. (I) has vulnerary and antiulcer activities and can be used to promote epithelial cell migration. (I) is useful for treating a wound and enhancing epithelialisation of a skin surface. The wound is especially skin wound including acute lesion such as traumatic lesion, burn, or surgical incision, chronic lesion such as chronic venous ulcer, diabetic ulcer or compression ulcer and the wound is further monitored for closure. The telomerase activity or TERT expression is increased in epithelial cells at the site of treatment and also in fibroblasts or endothelial cells at the site of treatment. The epithelial cells are especially keratinocytes. A polynucleotide encoding TERT is useful for the preparation of a medicament for treatment of a wound or an epithelial surface in a human or animal. An epithelial cell with increased telomerase activity or increased expression of TERT is useful for preparation of a medicament for the treatment of a wound in a human or animal. (I) is also useful for treating wounds of other epidermal surfaces including mucosal surfaces such as bronchus, mouth, nose, oesophagus, stomach or intestine. The present sequence encodes human TERT (hTERT), which is given in the exemplification of the present invention. hTERT is located to chromosome 5.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Query Match	45.8%;	Score 1600.8;	DB 25;	Length 4015;
Best Local Similarity	68.9%;	Pred. No. 0;		
Matches 2379;	Conservative 0;	Mismatches 967;	Indels 108;	Gaps 9
QY	23	CCCGCGCTTGAGCAATGACCCGCGGCTCCTCGCTTCCCGCGGCTGCGCTCTCTGCTCGG	82	
Db	40	CCCGGCAACCCCGGATGCGGCGCTCCCGCTGCCAGCCGTCGCTCCCTGCTCGG	99	
QY	83	CAGCCGATACCCGAGAGTGTGGCCGTGTGGCAACTTTTGTGCGGCGCTGTGGGCCCCGAGGG	142	
Db	100	CAGCCATACCGCGAGGTGCTGCCGTGTGGCACGTTCTGTGCGGCGCTCTGGGCGCCGAGG	159	
QY	143	CAGGCGGCTGTGTCAACCCGGGACCCGGAAGATCTACCGCACTTTGGTGTGCCAATGCT	202	
Db	160	CTGGCGCTGTGTGACGCGCGGGGACCCGCGCGCTTTCGCGCGCTGTGGCCAGTGCT	219	
QY	203	AGTGTGATGCACTGGGGCTCAGAGCTCCACTGCGGACCTTTCTCTTCACACGAGTGTC	262	
Db	220	GGTGTGCTGCTCGGACGACGCGCGCGCCCGCGCCCTCTCTTCGCGCAGGTGTC	279	
QY	263	ATCCCTGAAGAGCTGTGTGCGCAGGGTGTGCAGAGACTCTGCGAGCGCAACGAGAGAA	322	
Db	280	CTGCTGAAGAGCTGTGTGCCCGAGTGCTGACAGGCTGTGCGAGCGCGGCGCGAAGAA	339	
QY	323	CGTGTGCTTTTGGCTTTGAGTGTCTTAAACGAGGCCAGAGCGGGCTTCCCATGGCTTT	382	
Db	340	CGTGTGCTTTGGCTTTCGCGCTGTGTGACGGGGCCCGCGGGGGCCCCCGCGAGGCTTT	399	
QY	383	CACATGAGCTGGGTAGCTACTTGGCCCAACATGTTATTGAGACCTGCGTGTCAAGTGG	442	
Db	400	CACCAACGAGGTGGCGAGCTACTCTGCCCAACACGGTGAACCGACGCACTCGGGGGAGCGG	459	
QY	443	TGCATGATGTACTCTTTGAGCCGAGTGGGCGCAGCACTGTGCTTACCTGCTGCACAA	502	
Db	460	GGCTGTGGGCTGTGTGTGCGCGCGTGGCGACGACGTGTGTGTCACTCTGGCAGG	519	

503 CTGTGCTCTTTATCTTCTGTGTCCTCCAGCTGTGCTACAGGTGTGTGGTCTCCCT 562  
520 CTGGCGCTCTTTGTGTGTGGTCTCCAGCTCGCTATACAGGTGTGGCGCTCCGCT 579  
563 GTACCAAAATTGTCCACACAGGATATCTGGCCCTCTGTGTCCGCTAGTTTACAGGCCAC 622  
580 GTACCAAGTCTGGCGCTCGCACTCAGGCCCGCCCGCCACACAGCTAGTGTGACCCGGAAG 639  
623 CCGACCCGTGGGAGGAATTTCACTAACCTTAGGTTCTTACACAGATCAAGAGCAGTAG 682  
640 CGCTCTGGATGC-----GACCGGCCCTGGAACCATAG 672  
683 TCGCCAGAGACACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 CGTCAGGAGGCCGGGTCCCTTCGGCTCTCCAGCCCGGTGCGAGAGGCGGGGG 732  
743 TCTCACAGTACAAGTGTGCTTCAGTAAAGAGGCGAGATGTATCTGTCCGAGAGT 802  
733 CAGTGCAGCGGAAGTCTGCGGTTGCCAAGAGGCCCGAGGCTGGCGCTGCCCTGAGCC 792  
803 GGAGAGGAGCCACACAGGAGGTGTCTACAAACCCATCAGGCAATCATGGTGCACAG 862  
793 GGAGGAGGCCCTTGGGAGGAGGCTCTGGGCCACCCGGGCGAGACGCGTGGACCGAG 852  
863 TCCTGCTCGTCCCGAGGTGCTACT-----GCAGAGAAAGATTGTGTCTCTAA 913  
853 TGACCGTGGTTCTGTGTGTGTCACTGCGCAGACCGCGGAAGACCATCTTTTGA 912  
914 AGGAAAGGTGTCTACCTGATCTCTC---TGGGTGGTGTGTGTAAACACAGCCAG 970  
913 GGGTGCCTCTCTGGCAGCGCCACTCCACCATCGTGGGCGCGCAGCACACGCGGG 972  
971 CTCACAT---CTGTGTGTCAACACCGCCGCAAAATGCCCTTCAGCTCAGGCCATTAT 1027  
973 CCCCCATCCATCGCGGCCACACACCTCCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGAGATGGCCAAAGCGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTACTCTCTCAGGCGA---CAAGGAGCAGTGGCGCTCTCT 1089  
1088 CTTACTCAGCACTCCAGCCTACTTGTACTTGGGCCAGGAGACTGGTGGAGATCATCT 1147  
1090 CTTACTCAGCTCTCTGAGGCCCGCTCTGAGCTGGAGGCTGTGGAGACCATCT 1149  
1148 TCTGGCTCAAGGCTAGGACATCAGGACCACTCTGAGGACACACGCTCTATCGCGTGC 1207  
1150 TCTGGCTTCCAGGCGCTGGATGCGCAGGCACTCCCGCAGGTTGCCCGCTGCGCCAGCG 1209  
1208 ATACTGGCAGATGGGCCCTGTTCACACAGCTGCTGTGAACCATGCGAGTGGCAATA 1267  
1210 CTACTGGCAATGGGCCCTGTCTTCTGAGCTGCTTGGGAACACCGCAGTGGCCCTTA 1269  
1268 TGTACAGCTCTCAGGTGCATATTGAGGTTTCGAAACAGAAACCAACAGGTGACATGC 1327  
1270 CGGGTGTCTCTCAAGAGCACTGCCCGCTCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGCC 1342  
1330 CTGTGCCGGGAGAACCCAGGCTCTGTGGGGCCCGCGAGGAGGAGACAGACCC 1389  
1343 ACCGACCTCATGATTTGTCTCCGCTGCAACAGCTCCCTGGCAGGTATATGGTTTTCT 1402  
1390 CGGTGCTGTGTGAGCTGTCTCCGCGAGCAGCAGCGCTCTGAGGTGTACGGCTTCT 1449  
1403 TCGGCCCTGTCTCTGCAAGGTGTGTCTGTCTGTCTGGGGTACAGGCACAATGAGCG 1462  
1450 GGGGCCCTGCTGCGCGCTGTGTGCCCGCCAGGCTCTGGGGCTTCCAGGCACAAACGAG 1509  
1463 CGCTTCTTTTAAAGAACTTAAAGAAAGTTTCTCTGTGGGGAATATCGCAAGCTATCACT 1522  
1510 CGCTTCTCAGGAACACCAAGAAAGTTTCTCTCTGTGGGGAAGCATGCCAGCTCTCGCT 1569

Qy 1523 GCAGGAATGATGTGGAGATGAAGTAGAGATTGCCACTGGCTCCGACAGCCCGGG 1582  
Db 1570 GCAGGAGCTGAGTGAAGATGAGCGTGGGACTTGCCTTGGCTGGCAGAGCCGAG 1629  
Qy 1583 GAAGGACCGTGTCCCGCTGACAGACACCGTCTGAGGAGAGGATCCTGGTACGTTCT 1642  
Db 1630 GGTGGCTGTTCGGCGCGCAGACACCGTCTGCTGAGGAGATCCTGGCAAGTTCT 1689  
Qy 1643 GTTCTGGCTGATGGACACATAGTGTGACAGCTGTAGTCACTTCTTTTACATCAGGA 1702  
Db 1690 GCACTGGCTGATGAGTGTGTAGTGTGAGTGTCTCAGGCTTCTTTTATGTCAACGA 1749  
Qy 1703 GAGCATTCCAGAAACACAGGCTCTTCTTCTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762  
Db 1750 GACCACTTCAAAGAAACAGGCTCTTTTCTACCGAAGAGTGTCTGAGCAAGTTGCA 1809  
Qy 1763 GAGCATTTGAGTTCAGCAACACCTTGTAGAGATGGGCTACGGAGCTGTACAGAGGA 1822  
Db 1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGGAGCTGTGGAAGCAGA 1869  
Qy 1823 GGTCAAGCATCACAGACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCC 1882  
Db 1870 GGTCAAGCAGCATCGGAAGCCAGGCCCGCTGTGACGTCAGACTCCGCTTCATCCC 1929  
Qy 1883 CAAGCCCAACGGCTTCCGCCCATTTGTGAACATGATGATATAGCATGGGTACAGAGCTTT 1942  
Db 1930 CAAGCTTACGGCTTCCGCCCATTTGTGAACATGACTACGTCGTGGGAGCCAGAACGTT 1989  
Qy 1943 GGCAGAGAGGAGCAGGCCACGATTTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCT 2002  
Db 1990 CCGCAGAGAAAGAGGCGCAGCGTCTCACTCGAGGTGAAGGACTGTTTCAGCGTCT 2049  
Qy 2003 CAATATGAGCGGACAAAACATCTCTACCTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
Db 2050 CAATAGAGCGGCGCGCGCCCGCCCTCTCTGGCGCTCTGTGTCTGGGCTGGAAGA 2109  
Qy 2063 CATCTACAGACTGCGCGGCTTGTGTGCTGGTGTGGTGTCTGAGCCAGACACCCAG 2122  
Db 2110 TATCCACAGGCTTGGCGCACCTTGTGTGCTGGGCGCCAGGACCCCGCCCTGA 2169  
Qy 2123 GATGTACTTTGTTAAGGAGATGTACCCGGGCTTATGATGCCATCCCCAGGGTAAGCT 2182  
Db 2170 GCTGTACTTTGTCAAGTGTGTGAGCGGCGCTACGACACCATCCCCAGGACAGCT 2229  
Qy 2183 GGTGAGGTTGTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
Db 2230 CACGGAGGTCACTCCAGCATCATCA---AACCCAGAAACAGTACTCGTGGCTCGGTA 2286  
Qy 2243 TGCAATGCTCCGAGAGATAGCCAGGCCAAGTCCAAGTCTCTTTAGGAGACAGGTCAAC 2302  
Db 2287 TGCGTGTCCAGAAAGCGCCCATAGGGCACTCCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
Qy 2303 CACCTCTCTGACTCCAGCATATAGTGGCGAGTTCCTTAAGCATCTGAGGATTCAGA 2362  
Db 2347 TACCTTGACAGACTCCAGCCGTACATGCAAGTTCGTGGCTCACTCTCAG-----GA 2400  
Qy 2363 TGGCAGTGCATCAGGAACTCCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
Db 2401 GACAGCCGCTCAGGAGTCCGCTGTCTATCAGCAGAGCTCTCTCTCTGAATGAGGCCAG 2460  
Qy 2423 CAGCAGCTGTTTGTACTTCTCTGCACTTCTCTGCTGCTCAAGTGTCTGTAAGATTTGTA 2482  
Db 2461 CAGTGGCTCTTCTGAGCTCTTCTACGCTTCATGTGCCACACCGCGTGGCATCAGGGG 2520  
Qy 2483 CAGGTGTATACAGTGGCAGGCACTCCCGAGGCTCCAGCTATCCACCTGCTCTG 2542  
Db 2521 CAAGTCTCTAGTCAAGTGGCAGGATCCCGAGGCTCTCACTCTCTCCAGCTGCTCTG 2580  
Qy 2543 CAGTCTGTGTTTCCGAGACATGAGAAACAAAGTGTGTTGCTGAGGTGCAAGCGGATGGGT 2602  
Db 2581 CAGCTGTGCTACGGCGACATGAGGAACAAAGCTGTTTGGGGGATTCGGCGGACCGGCT 2640  
Qy 2603 GCTTTTACGTTTTTGTGATGACTTCTCTGTGTGTGAGCGCTCACTTGGACCAAGCAAAAC 2662





503 CTGTGCTCTTTACTTTCTGTGTCGCCCCCAGCTGTGCTACAGGTGTGTGGTCTCCCT 562  
520 CTGCGCCTCTTTGTGTGTGTGCTCCAGCTCGCGCTTACCAAGGTGTGCGGCCCCCGCT 579  
563 GTACCAAAATTTGCCACCAACGATATCTGGCCCTCTGTGTCGCTAGTTACAGGCCAC 622  
580 GTACCAAGCTCGGCGCTGCACACTCAGGCCCGCCCCGACACGCTAGTGGACCCCGAAG 639  
623 CCGACCGTGGGAGGAAATTTCACTAACCTTAGGTCTTACAACAGATCAAGAGCAGTAG 682  
640 CGGTCTGGGATGC-----GAAACGGGCTTGAACCATAG 672  
683 TCGCCAGGAAGACCGAAACCCCTGCTTCATCTCAGAGGTACAAGAGAGCATCTGAG 742  
673 CGTCAGGAGGCGCGGGTCCCTTGGCCCTGCAGCCCGGGTGGAGGAGGCGGGGG 732  
743 TCTCACAGTACAGTGTGCTTCACTAAGAGGCCAGATGCTATCTCTTCCGAGAGT 802  
733 CAGTGCAGCCGAAGTCTGCGGTGCCCAGAGGCCAGGCGTGGCGTGCCTCTGAGCC 792  
803 GGAGGAGGACCCACAGGAGGCTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 GGAGCGACCCCGTTGGGACAGGGTCTTGGGCCACCCCGGCGACGAGCGGTGGACCGAG 852  
863 TCCTGTCTGCTCCCGAGTGCCTACT-----GCAGAGAAAGATTTGTCTTTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCACTGCCAGACCCGCGGAAGACCACTCTTTTGA 912  
914 AGAAAGGTCTGACTGAGTGTCTC---TGGGTGCGTGTCTTAACACACAGCCAG 970  
913 GGTGCGCTCTTGGACCGGCCACTCCACCCATCCGTGGGCCCGCAGCACCGCGGG 972  
971 CTCACAT---CTCTGTGTACACCCCGCCAAATGCTTTTACGCTCAGGCCATTTAT 1027  
973 CCCCCATCCACATCGCGCCACCACTCCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACGAGATTTCTTTTACTTCAGGGGAGATGGCAGAGCGCTTAACCCCTCAT 1087  
1033 CGAGACCAAGCATTTCTCTACTCTCTCAGGCGA---CAAGGAGAGCTGCGGCCCTCT 1089  
1088 CTTACTCAGCAACTCCAGCTTAACCTGAGTGGGCGCAGGAGACTGGTGGAGATCATCT 1147  
1090 CTTACTCAGCTCTTGAGGCCAGCTGACTGGCGTTCGAGGCTCGTGGAGACCATCT 1149  
1148 TGTGGGCTCAAGGCTAGGACATCAGGACCACTCTGACGAGACACACCGTCTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCCCTTGAATGCGAGGAGCTCCCGCAGGTTGCCCGCCTGCCCCAGCG 1209  
1208 ATACTGGCAGATGGGCCCTGTGTTCCAAAGCTGCTGGTGAACCATGACAGAGTGCATA 1267  
1210 CTACTGCAATGCGGCCCTGTTTCTGAGCTGTCTTGGAAACCGCAGCTGCCCCCTA 1269  
1268 TGTCAAGTCTCAGGTTCATTTGCAAGTTTGAACAGAAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGTGGGTCAGCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACAGCCC 1342  
1330 CTGTGCCCCGAGAGCCCCAGGGCTGTGTGGGCGCCCCCGAGGAGGAGACACAGACCC 1389  
1343 ACCGCACTCATGATTTGTCTCGGCTTGACAGCAGTCCCTGCGAGGTATATGTTTTCT 1402  
1390 CCGTGGCTGTGTCAGTGTCTCGCCAGCAGCAGCAGCCCTTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGGCTGTCTGCAAGGTGTGTCTGTAGTCTCTGGGATCCAGGCACAATGAGCG 1462  
1450 GCGGGCTGTCTGCGCGGTGTGTGCCCCAGGCTCTGGGGCTCCAGGCAACAGAAAG 1509  
1463 CCGCTCTTTAAGAACTTAAAGAGTTTCATCTGTTGGGGAATACGGAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAAACCAAGAAAGTTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGGAATGATGTGGAAGATGAAGATGAGGATTTGCCACTGGCTCCGAGCAGCCCGG 1582

1570 GCAGGAGCTGACGCTGGAAAGATGAGCTGCGGAGCTGCGCTTGGCTGCGAGAGCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGACACCGTCTGAGGAGAGATCTCTGGCTACGTTCTCT 1642  
1630 GGTGGCTGTGTTCGGGCGGAGACACCGTCTGCGTGGAGAGATCTCTGGCCAAAGTTCTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGGTACGCTGTTAGGTCACTTTTACATCACAGA 1702  
1690 GACTGGCTGATGAGTGTGCTGCTGAGCTGCTCAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCACTTCCAGAAAGAACAGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACCAGTTTCAAAGAAAGAGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATGGAGTCAAGCAACACCTTGAAGAGTGGCTGAGAGGCTGTCGGAAGCAGA 1822  
1810 AAGCATGTGAATCAGACAGCACTTGAAGAGGTGAGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTCAAGGATCACACAGGACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCATCCC 1882  
1870 GGTCAAGGAGCATCGGGAAGCCAGGCCCGCTGCTGAGTCCAGACTCCGCTTCATCCC 1929  
1883 CAAAGCCAAACGCTTGGGCGCAATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAAGCTGACGGGCTGCGGCCGATTTGTGAACATGAGCTACGCTGCGGAGCCAGAGCTT 1989  
1943 GGGCAAGGAAGAGCAGGCCAGCATTTTCAACAGCGCTCTCAAGACTCTTTCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGGAGCGTCTCACTCGAGGTTGAAGGCACTGTTTCAGCGTGTCT 2049  
2003 CAACTATAGCGGACAAACATCTCACTTATGGGTCTTTCTGTATCTGGGTATGAATGA 2062  
2050 CAACTACAGCGGCGCGCGCTCTCTGGGCGCTCTCTGTGTGGCTTGGGCTGAGCA 2109  
2063 CATCTACAGGACTGCGGGCGCTTGTCTGCTGCTGCTGCTGAGCAGACAGCAGCCAG 2122  
2110 TATCCACAGGCTGGCGACCTTGTGCTGCTGCTGCTGGGCGCCAGGACCCGCGCTGA 2169  
2123 GATGTACTTTGTTAAGGAGATGTGACCGGGGCTTATGATGCCATCCCCAGGGTAAAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGGATGTGACGGGCGCTGACGACACCATCCCCAGGACAGCT 2229  
2183 GGTGGAGTGTGTTCCAAATATGATCAGGCACTCGGAGAGCAGCTACTGTATCCGCCAGTA 2242  
2230 CACGGAGTCTACGCGACGATCATCA---AAGCCAGAAACAGTACTGCGTGGCTCGGTA 2286  
2243 TGCAGTGGTCCGAGAGATAGCCAGGCGCAAGTCCCAAGTCTTTAGGAGACAGGTCA 2302  
2287 TSCCGTGGTCCAGAGGCGGCCCATGGGACGTCGCGAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGCTACATGCGACAGTTCTGGTCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCATGAGGAACTCCGTTGTCTATCAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCGCTGAGGAGTGCCTGCTCATCGAGCAGGCTCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGAATTTCTTCTGCACTTCTCTGGTACAGTGTCTGTAAGATTGTTGA 2482  
2461 CAGTGGCTCTTCGACGCTCTTCTACGCTTCTATGCGCACCGCCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATCGCAGTCCAGGGCATCCCCAGGGCTCCAGCTATCCACCCCTGCTG 2542  
2521 CAACTCTACGTCCTGAGTGCAGGGATCCCGAGGCTCCATCTCTTCCAGCTGCTCTG 2580  
2543 CAGCTCTGTTTTCGAGACATGAGAAACAAAGTGTGTTGCTGAGGTGACGCGGGATGGGTT 2602  
2581 CAGCTGTCTTACGCGACATGAGAAACAAAGTGTGTTGCGGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGTATGACTTTCTGTGTGGTGAAGCTCCTGAGTGGACAGCAAAAAC 2662

2641 GCTCTCGGCTTGGTGGATGATTTCTTGTGTGACACTCACCACCCAGCGAAAC 2700  
2663 CTTCTCAGACACCTCGTCAATGGCGTTCCTGAGTATGGTGCATGATAAACTTGCAGAA 2722  
2701 CTTCTCAGACACCTCGTCCGAGGTTCCTCGATGATGCTCGTGGTGAATCTTCCGAA 2760  
2723 GACAGTGGTGAATCTCCCTGTGGAGCTGATACCTCGGTGTGAGCTCCATACCAAGCT 2782  
2761 GACAGTGGTGAATCTCCCTGTAGAGACAGGCTTGGTGTGACGCTTTTGTTCAGAT 2820  
2783 GCTGTCTCACTGCTCTTCCCTGTGGTGTGCTGTGTCGACACTCAGACTTTGGAGGT 2842  
2821 GCGGGCCACGGGCTATTCCTGTGTGGCGCTGCTGTGTGATACCGGACCCCTGGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTATGCGCCAGACCTCAATTAAGACGACCTCAGCTCCAGAG 2902  
2881 GCGAGGACTACTCAGGTATGCGCGGACCTCCATCAGAGCCAGCTCAGCTTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAAGACCATGCGGAACAAGCTCCTGTGCGTCTTTCGGTGTGAAGTG 2962  
2941 CGCTTCAAGCTGGGAGGAACATGCTCGAACTCTTTTGGGTCTTTCGGCTGAAGTG 3000  
2963 TCAGGTCTATTTCTAGACTTGAGTGAAGACCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCAGGCTGTTTCTGGATTTGAGGTGAACAGCCCTCCAGACGGGTGTGCACCAATCTA 3060  
3023 CAAGATCTTCTCTTTCAGGCTTACAGGTTCATGATGATGATTCAGTTCCTCTTGA 3082  
3061 CAAGATCTTCTCTGCGAGGCTACAGGTTTACGATGTGTGCTGAGTCCCATTTCA 3120  
3083 CAGGCTGTAGGAAGACCTCAATCTTCTTGGGATCATCTCCAGCCAGCATCTTG 3142  
3121 TCAGCAAGTTTGAAGAAGCCCAATTTTCTGCGGCTCATCTGTACAGGCTCCT 3180  
3143 CTGCTGTCTATCTTCAAGTCAAGATCCAGATGACACTAAG-----GCCTC 3193  
3181 CTGCTACTCATCTTGAAGCCAGGATCCAGGATGTGCTGGGGCCAGGGCGCGC 3240  
3194 TGCTCTTCTCTTCAAGCCGACATTTGGTCTGCTACAGGCTTCTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTCTCGAGGCTGCGAGTGGTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
3254 GGCTGTCTATCTGCTATCAAAATGCTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313  
3301 GACTCGACACCGTGTCACTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGCTGAGCTGACCCAGC 3373  
3361 GCTGAGTGGGAAGCTCCCGGGACGACGCTGACTGCTGAGGCGCGAGCAACCCGGC 3420  
3374 CTAAGCAGACTTTCAGACCACTTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCATCTCGACTGA 3454

JLT 11

50320

AAV60320 standard; cDNA; 4023 BP.

AAV60320;

04-DEC-1998 (first entry)

Human telomerase gene referred to as hEST2.

Catalytic subunit; human; telomerase; telomere maintenance; diagnosis; treatment; cancer; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 59...3458

/\*tag= a

XX WO9837181-A2.  
XX 27-AUG-1998.  
XX 20-FEB-1998; 98WO-US03404.  
XX 30-OCT-1997; 97US-0064322.  
XX 20-FEB-1997; 97US-0038750.  
XX 20-MAY-1997; 97US-0047151.  
XX 01-AUG-1997; 97US-0054549.  
XX 14-AUG-1997; 97US-0055762.  
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
XX Counter CM, Meyerson M, Weinberg RA;  
XX P-P8DB; AAW71376.  
XX WPI; 1998-495367/42.  
XX DR DR DR  
XX PT develop products for increasing or reducing the life span of cells  
XX PT such as cancer cells or transformed cells  
XX PS Claim 5; Fig 5A-B; 96pp; English.  
XX CC The present sequence encodes the catalytic subunit of a human  
XX CC telomerase holoenzyme. Disruption of the telomerase gene alters  
XX CC telomere maintenance. The DNA is essential for telomerase activity,  
XX CC and the protein is physically associated with telomerase and a  
XX CC constituent of active telomerase complex. The products can be used  
XX CC for increasing or reducing the lifespan of cells such as cancer cells  
XX CC or transformed cells. They can also be used in the diagnosis and  
XX CC treatment of malignancies. In addition, cells with a longer lifespan  
XX CC can be transplanted into or grafted onto an individual (e.g. as skin  
XX CC grafts, as systems for delivery of therapeutic proteins, such as hormones  
XX CC and enzymes), to whom they provide therapeutic benefit.

XX SQ Sequence 4023 BP; 668 A; 1363 C; 1277 G; 715 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 19; Length 4023;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGCGCTTGAAGCAATATGACCCGCGTCTCTGTTGCCCGCGGTGCGTCTCTGTCGG 82  
Db 43 CCGCGCCACCCCGCGATGCGCGCGCTCCCGCTGCGAGCGCTCCCTGCTGG 102  
QY 83 CAGCCGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 142  
Db 103 CAGCCACTACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGGCGCTGGGCGCCGAGGG 162  
QY 143 CAGCGGCTTGTGCAACCCGCGGACCCGAGATCTACCGACATTTGTTGCCCAATGCTT 202  
Db 163 CTGGCGGCTGTGTCAGCGCGGGGACCCGCGGCTTTCCGCGCGCTGTGTGCGCGCTT 222  
QY 203 AGTGTGATGACTGCGGGCTCAGACCTCCAGCTGCGGACCTTTCTTCCACCGAGTGC 262  
Db 223 GGTGTGCTGCTTGGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 282  
QY 263 ATCCCTGAAGAGCTGTGTGGCCAGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 283 CTGCTGAAGAGCTGTGTGGCCCGCGAGTGTGACAGGCTGTGCGAGCGCGCGCGAGAA 342  
QY 323 CGTGCTGCTTTTGGCTTTGAGCTGTATTACGAGCGGCGGCGGCGCTCCCATGCTT 382  
Db 343 CBTGCTGCTTTCGCTTTCGCGCTGCTGAGCGGGGCGCGGGGGCGGGGGCGGGGGCTT 402  
QY 383 CACTAGTAGCGTGTAGCTACTTGTGCCCAACACATGTTATTGAGACCTTGGGTGTCACTGG 442  
Db 403 CACCACCGGTGCGCACTACTTGTGCCCAACACACGTTGACCGCACTGCGGGGAGCGG 462  
QY 443 TGCATGATGTACTGTTGAGCGGAGTGGGCGGACGACTGCTGTGTACTCTGTCTGCACA 502



2603	GCTTTTACGTTTGTGATGACTTCTCTGTGTGACGCTCACTTGGACCAAGCAAAAC	2662	OS	Homo sapiens.
2644	GCTCTGCGTTTGTGTGATGATTTCTTGTGTGACACCTCACCTCACCCACGGAAC	2703	XX	WO9935243-A2.
2663	CTTCTCAGACACCTGTCCTGAGTATGGTGCATGATTAACCTTGCAGAA	2722	XX	15-JUL-1999.
2704	CTTCTCAGACACCTGTCCTGAGTATGGTGCATGATTAACCTTGCAGAA	2763	XX	12-JAN-1999; 99WO-US00682.
2723	GACAGTGGTGAATCTCCCTGTGAGCGCTGCTACCTGGTGTGAGCTCCATACCAAGCT	2782	XX	21-APR-1998; 98US-0063657.
2764	GACAGTGGTGAATCTCCCTGTGAGCGCTGCTACCTGGTGTGAGCTCCATACCAAGCT	2823	XX	12-JAN-1998; 98US-0071220.
2783	GCTGTCTACTGCTGTTTCCCTGTGAGCGCTGCTACCTGGTGTGAGCTCCATACCAAGCT	2842	XX	13-JAN-1998; 98US-0071455.
2824	GCGGCGCCACGCGCTATTCCTGTGAGCGCTGCTACCTGGTGTGAGCTCCATACCAAGCT	2883	XX	(COLD-) COLD SPRING HARBOR LAB.
2843	GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTTCACTTCCAGAG	2902	XX	Beach DH, Hamon GJ, Wang J;
2884	GCAGAGGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTTCACTTCCAGAG	2943	XX	WPI; 1999-444196/37.
2903	TGCTTTCAAGCTTGGGAAGCAATCGGGAACAGCTTCTGCTGGTCTTGGCGCTGAAGTG	2962	XX	P-PSDB; AAY28401.
2944	CGGCTTCAAGCTTGGGAAGCAATCGGGAACAGCTTCTGCTGGTCTTGGCGCTGAAGTG	3003	XX	Increasing proliferative capacity of cells useful for promoting
2963	TCAGGCTTATTTCTAGACTTGCAGTGAACAGCTTCCAGACAGTCTGCATCAATATATA	3022	XX	wound healing
3004	TCAGGCTTATTTCTAGACTTGCAGTGAACAGCTTCCAGACAGTCTGCATCAATATATA	3063	XX	Claim 4; Page 65-70; 73pp; English.
3023	CAAGATCTTCTGCTTCCAGCTTACAGTTCATGATGATGATGATGATGATGATGATGAT	3082	XX	This sequence encodes the human EST2 protein, and can be used in the
3064	CAAGATCTTCTGCTTCCAGCTTACAGTTCATGATGATGATGATGATGATGATGATGAT	3123	XX	method of the invention. The method is for increasing the proliferative
3083	CCAGCGTGTAGGAAGCACTCAGTTCATGATGATGATGATGATGATGATGATGATGAT	3142	XX	capacity of cells, and comprises contacting the cell with a
3124	TCAGCAAGTTTGAAGAACCCCACTTTTCTGCGCGTCACTCTGACAGCGCTCCCT	3183	XX	telomerase-activating therapeutic agent (TATA). The method can be used
3143	CTGTGATGCTATCTGAGGTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG	3193	XX	for extending the life-span of cells, e.g. by increasing the number of
3184	CTGTGATGCTATCTGAGGTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG	3243	XX	mitotic divisions. They can be used for e.g. the extension of skin or
3194	TGGTCTTCTTCTTCCAGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3253	XX	other epithelial cell cultures or grafts, the expansion of mesenchymal
3244	CGGCGCTTCTGCTTCCAGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3303	XX	cell cultures or grafts, and the expansion of chondrocyte or osteocyte
3254	GGCTGCTTCTGCTTCCAGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3313	XX	cultures or grafts. They can be applied to e.g. neuronal, haematopoietic,
3304	GACTCGACCGTGTACCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3363	XX	epithelial, pancreatic, hepatic, chondrocytic and osteocytic stem and
3314	GCTGTGCGGAAGCTTCCAGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3373	XX	progenitor cells in vivo, in vitro or ex vivo protocols. The methods
3364	GCTGTGCGGAAGCTTCCAGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	3423	XX	can be used for promoting the healing of wounds resulting from
3374	CCTAAGCAGACACTTTCAGACCACTTTTGGACTAA 3407		XX	e.g. surgery, burns, inflammation or irritation or ulcers resulting from
3424	ACTGCCCTCAGACTTTCAGACCACTTTCGGACTGA 3457		XX	e.g. venous disease (venous stasis ulcers), excessive pressure (decubitus
ULT 12			XX	ulcers) or arterial ulcers. They can also be used to enhance tissue
89424			XX	regeneration processes, e.g. of the skin, hair and/or fingernails. They
AAx89424 standard; cDNA; 4027 BP.			XX	can also be used for treating age-related conditions, e.g. atrophy of the
AAx89424;			XX	skin through loss of extracellular matrix homeostasis in dermal
22-SEP-1999 (first entry)			XX	fibroblasts, age-related macular degeneration caused by accumulation of
Human EST2 coding sequence.			XX	lipofuscin and downregulation of a neuronal survival factor in retinal
EST2; proliferative capacity; cellular proliferation; decubitus ulcer;			XX	pigmented epithelial (RPE) cells, and atherosclerosis caused by loss of
telomerase-activating therapeutic agent; cell life-span extension;			XX	proliferative capacity and overexpression of hypertensive and thrombotic
venous disease; venous stasis ulcer; excessive pressure; arterial ulcer;			XX	factors in endothelial cells. Expanded populations of normal or
tissue regeneration enhancer; atherosclerosis; therapy; ss.			XX	genetically engineered rejuvenated cells could be used for autologous or

Query Match 45.8%; Score 1600.8; DB 20; Length 4027;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
SQ Sequence 4027 BP; 674 A; 1361 C; 1277 G; 715 T; 0 other;  
QY 23 CCAGCGCTTGTGACCAATGACCCGCGCTCTCTGTTGCCCGCGTGGCTCTCTGCTGCG 82  
DB 41 CCAGCGCACCCCGGATGCGCGGCTCCCGCTGCGAGCCGCTCCCTGCTGCG 100  
QY 83 CAGCGGATACCGGAGGTGTGCGCGCTGCGCAACCTTTGTGCGCGCTTGGGGCCCGAGGG 142  
DB 101 CAGCGGATACCGGAGGTGTGCGCGCTGCGCAACCTTTGTGCGCGCTTGGGGCCCGAGGG 160  
QY 143 CAGCGGCTTGTGCAACCCCGGAGCCCGCAAGATCTACCGCACTTTGTTGCTCCCAATGCT 202  
DB 161 CTGGCGCTTGTGCAACCCCGGAGCCCGCAAGATCTACCGCACTTTGTTGCTCCCAATGCT 220  
QY 203 AGTGTGATGCACTGGGGCTCAGCGCTTCCAGCTTCCAGCTTCCCTTCCACAGGCTC 262  
DB 221 GGTGTGCTGCTGCGGACG 280

```
263 ATCCCTGAAAGAGCTGTGGCCAGGGTTGTGAGAGACTCTCGGAGCGCAAGAGAGAAA 322
281 CTGCTCTGAGAGAGCTGTGGCCAGTGTCTGAGAGCTGTGCGAGCGGGCGCAAGAAA 340
323 CGTCTGGCTTTTGGCTTTTGGCTTTTAAAGAGGCGAGAGCGGGCTTCCCATGGCCCTT 382
341 CGTCTGGCTTTGGCTTTTGGCTTTTAAAGAGGCGAGAGCGGGCGGGCGGGCGGGCTT 400
383 CACTAGTAGGCTGGCTAGCTACTTGGCCACACTGTTATTGAGACCTGCGTGTGAGTGG 442
401 CACACAGCGTGGCGAGCTACTTGGCCACACTGTTAAAGAGGCGAGAGCGGGCGGG 460
443 TGCATGATGCTACTTGTGAGCGGAGTGGGCGAGCACTGCTGCTTACCTGCTGGCACA 502
461 GGGCTGGGGCTGCTGTTGGCGCGCTGGGCGAGCAGTGTGTTTACCTGCTGGCAG 520
503 CTGTGCTTTTATCTTCTGGTGGCCCCGAGCTGTGCTTACAGAGTGTGTGGTCTCCCT 562
521 CTGGCGCTCTTTGTGTGGTGGCTCCGAGCTGGGCTTACCAAGTGTGGGGCGCGCT 580
563 GTACCAAAATTTGTGCCACACAGCATATCTGGGCGCTCTGTGCTCCGCTAGTTACAGGCCAC 622
581 GTACCAAGCTGGCGCTGCTCAGGCGCGGCGCGCGCACACGCTAGTGTGACCCCGAG 640
623 CGACCGGTGGGAGGAAATTTCACTAACCTTAGTGTCTTAAACAGATCAAGAGCAGTAG 682
641 GCGTCTGGGATGC-----GAAACGGGCGCTGGAACCATAG 673
683 TCGCCAGGAAGACCGAAACCCCTGGCTTGCATCTCGAGGTCAAAAGAGCATCTGAG 742
674 CGTCAGGGGCGCGGGTCCCTCTGGCTGCGAGCGCGGGTGCAGAGGCGGGGG 733
743 TCTCAACAGTAAAGTGTGCTTACGCTAAGAGGCGAGTCTATCTGCTCCGAGAGT 802
734 CAGTGGCAGCGGAAGTCTGCGTGTGCCAAGAGGCGCAGGCGTGGCGCTGCCCTGAGCC 793
803 GGAGGGGACCCACAGCAGCTGTCTACCAACCCCATCAGCAATCATGGTGGCCAG 862
794 GGAGCGGACGCGCTTGGGAGGGGCTTGGGCGCACCCCGGCGAGGACGCGTGAGCCGAG 853
863 TCTGTCTGGTCCCCGAGTGGCTACT-----GCAGAGAAAGATTGTCTTCTAA 913
854 TGACCGTGGTTCGTGTGGTGTCACTGCCAGACCGCGCGAAGAACCACTCTTTTGA 913
914 AGGAAAGGTCTTGACCTGAGTCTCTC---TGGTGGTGTGCTGTAAACAGAGCCAG 970
914 GGTGGCTCTCTGGCAGCGGCACTTCCACCCATCCGTTGGGCGCCAGCACGCGGG 973
971 CTCACAT---CTCTGCTGTACACCCCGCAAAATGCTTTTCACTCAGCTCAGGCCATTTAT 1027
974 CCCCCATCCACATCGCGGCCACCACTCCCTGGGACAGCGCTTGTCCCGGTGTACGC 1033
1028 TGAGACGAGCATTTCTTTTACTCAGGGGAGATGGCCAGAGCTCTAAACCCCTCATTT 1087
1034 CGAGACCAAGCACTTCTCTACTCTCTCAGGCA---CAAGGAGCAGCTGCGGGCTCTCTT 1090
1088 CTTACTCAGCAACCTCCAGCTTAACCTTGACTGGGCGCAGGAGCTGTTGGAGATCATCTT 1147
1091 CCTACTCAGCTCTCTAGGCGCCAGCTGACTGGCGCTCGGAGGCTGTGGAGACCATCTT 1150
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACACCTCTATCGGCTG 1207
1151 TCTGGGTTCCAGGCTTGTAGTGCAGGGGACTCCCGCGAGGTTTGGCCCGCTGCCCCAGCG 1210
1208 ATACTGAGATCGGCCCTCTGTTCCAACTGCTGTGGTGAACCATGAGAGTGCACATA 1267
1211 CTACTGGCAATGCGGCCCTCTGTTCTGAGTGTGTTGGAAACCAAGCGAGTGGCCCTTA 1270
1268 TGTCAAGCTCTCAGGTCATTCAGAGTTTCCAAACAGCAACCAACAGGTGACAGATGC 1327
1271 CGGGGTGCTCTCTCAAGACGCACTGCGCGCTGGAGCTGGGCTCACTCTATGATGAGAGCAG 1330
1328 CT-----TGAACACCAAGCC 1342
1331 CTGTGCCCGGAGAAAGCCCGAGGCTCTGTGGCGGCCCCCGAGAGGAGGACACAGACCC 1390
1343 ACCGACCTCATGGAATTTGCTCCGCTCTGCACAGAGTCCCTGGGAGGTATATGTTTCT 1402
1391 CGGTCCGCTGGTGCAGTGTCTCCGCGACACAGCAGCCCTGGCAGGTGTACGCTTCGT 1450
1403 TGGGGCTGTCTCTGAAAGTGTGTCTGTAGTGTCTCTGGGGTACCGGCAATGAGCG 1462
1451 GGGGGCTGCTCTGCGCCGCTGTGTCGCCCGAGGCTCTGGGGCTCCAGGCACACGAAAG 1510
1463 CCGCTCTTTTAAAGAACTTAAAGAACTTCACTCTGTTGGGAAATACGCGAAGCTATCACT 1522
1511 CCGCTTCTCAGNAACACNAGAGTTCACTCTCCTGGGAAAGCATGCCAAGCTCTGCT 1570
1523 GAGGAACTGATGTGGAAGATGAAAGTAGAGATGTCACATGCGCTCCGAGCAGCCCGGG 1582
1571 GAGGAGCTGAGTGTGGAAGATGAGCTGCGGGGCTGGGCTTGGCTGGCGAGGAGCCAGG 1630
1583 GAAGGACCGTGTCCCGCTGCAGAGCAGCTCTGAGGAGAGGATCTGGCTAGCTTCCT 1642
1631 GGTGGCTGTGTCCGGCGCGAGAGCACCTCTGCTGAGAGATCTGGGCAAGTTCCT 1690
1643 GTTCTGGCTGATGAGACACATACGTGGTACAGCTCTTAGGTCACTTTTACATCACAGA 1702
1691 GCACTGGCTGATGAGTGTGTACGTCTGAGCTCTCAGGTCTTTCTTTTATGTACGGA 1750
1703 GAGCACTTCCAGAGAAACAGCTCTTCTTACCGTAAGAGTGTGGAGCAGCTGCA 1762
1751 GACCAAGTTTCAAAGAAACAGCTCTTTTCTACCGAAGTGTCTGGAGCAAGTTGCA 1810
1763 GAGCACTGAGTTCAGGCAACACCTTGAAGAGTGGCTACGGGAGCTGTCAACAAGAGA 1822
1811 AAGCATTTGAAATCAGACAGCACTTGAAGAGGCTCAGCTCGGAGAGCTGTGGAAGAGA 1870
1823 GGTCAAGATCAGCAGGACACCTGGCTAGCATGCCATCTGCAGACTGGCTTCATCCC 1882
1871 GGTCAAGCAGCATCGGGAAGCCAGGCGCGCTCTGACGTCACAGCTCTCAGCTTCATCCC 1930
1883 CAAGCCAAACCGCTGCGGGCCCATTTGTGAACATGAGTTATAGCATGGGTACCAAGAGCTT 1942
1931 CAAGCTGACGGCTGCGGCGGATTGTGAACATGCACTACGTCTGGGAGCCAGAACGTT 1990
1943 GGGCAGAGGAAGCAGGCGCCAGCATTTACCCAGCGTCTCAAGACTCTCTTCAAGCATGCT 2002
1991 CGGCAGAGAAAGAGGCGCGAGCTCTCACTCGAGGTTGAAGGCACTGTTCAGCGTCT 2050
2003 CAACATATGAGCGGACAAACATCTCACTTATGGGCTCTTGTACTGGGTATGAATGA 2062
2051 CAACATCAGCGGGCGCGGCGCCCGGCTCTTGGGCGCTCTGTGTGGGCTGTGACGA 2110
2063 CATCTACAGGACTGGCGGGCTTTGTGTGCTGGTGGTGGTCTGGAACAGACACCCAG 2122
2111 TATCCACAGGCGCTGGCGCACTTCTGCTGCTGCTGCTGGGCGCTCTGTGTGGGCTGTGACGA 2170
2123 GATGTACTTTTGAAGCAGATGTGACCGGGGCTATGATGCCATCCCCCAGGGTAACT 2182
2171 GCTGTACTTTGTCAAGTGTGATGTGACGGGCGGTACACACATCCCCCAGGAGGCT 2230
2183 GGTGAGGTTGTGTCAAATATGATCAGGCACTCGGAGAGCACTGTGTATCCGCCAGTA 2242
2231 CACGAGGTCATCGCCAGCATCATCA---AACCCAGAACACGCTACTGCTGGTGGTTCGTA 2287
2243 TGCAGTGTTCGGAGAGATAGCCAGGCCAGTCCCAAGTCTCTTTTAGGAGACAGGTTCAC 2302
2288 TGCCGTGTTCAGAAAGCGCGCCATGGGACGTCGCAAGGCTTCCAGAGGCTTCAAGAGCAGCTTC 2347
2303 CACCTCTCTGAGCTTCCAGCCATACATGGGCGAGTTCCTTTAAGCATCTGCAGGATTCAGA 2362
2348 TACCTTGACAGACTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2401
2363 TGCCAGTGCATGAGGAACTCCGTTGTATCTGAGCAGAGCATCTCTATGATGAGAGCAG 2422
```

2402 GACGAGCCGCTGAGGATGCGTCGTCATCGACGAGCTCTCCCTGAATGAGCCAG 2461  
|||||  
2423 CAGCAGCTGTTGACTTCTCTCTGCACTCTCTGCGTCACAGTGTGTAAGATTGGTGA 2482  
|||||  
2462 CAGTGGCTCTTCGACGCTCTCTACGCTTATGTGCACACGCGTGGCATAGGGG 2521  
|||||  
2483 CAGTGTCTATACGAGTCCAGGGCATCCCGAGGGTCCAGCTATCCACCTGCTCTG 2542  
|||||  
2522 CAACTCTACGTCAGTCCAGTCCAGGGGATCCCGAGGGTCCATCTCTCCACGCTGCTG 2581  
|||||  
2543 CAGTCTGTGTTGGAGACATGAGAACAGCTCTTTGCTGAGTGCAGCGGATGGGTT 2602  
|||||  
2582 CAGCCTGTGCTACGCGCATGAGAACAGCTGTTTGGGGGATTCGGCGGACGGCT 2641  
|||||  
2603 GCTTTTACGTTTGTGTGACTTTCTGTTGGTGACGCTCTCACTTGACCAACAAAC 2662  
|||||  
2642 GCTCTCGGTTTGGTGGATGATTCTTGTGTGACACCTCACCTCACCCACGCAAAAC 2701  
|||||  
2663 CTTCTCAGCACCTTGTGCTTCCATGGGTTCTGAGTATGGTGCATGATAAATTGAGAA 2722  
|||||  
2702 CTTCTCAGCACCTTGTGCTTCCAGTGTCTGAGTATGGTGCATGATAAATTGAGAA 2761  
|||||  
2723 GACAGTGTGAACTTCCCTGTGAGCCTGTGACCTGGTGTGCTGCTCATACAGCT 2782  
|||||  
2762 GACAGTGTGAACTTCCCTGTGAGACGAGGCGCTGGTGGCAGCGCTTTGTTCAGAT 2821  
|||||  
2783 GCTGCTCAGCTGTTTCCCTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
|||||  
2822 GCGGCGCCACGCGCTATTCCCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2881  
|||||  
2843 GTTCTGTGCTACTCAGGTTATGCGCAGACCTCAATTAAGACGAGCTCACCTTCCAGAG 2902  
|||||  
2882 GAGAGCGACTACTCAGCTATGCGCGACCTCTCCATCAGACCGACTCTACCTTCAACG 2941  
|||||  
2903 TGTCTTCAAAGTGGGAGAACCATCGGAAACAGCTCTGCTGCTGCTGCTGCTGCTGCTG 2962  
|||||  
2942 CGGCTTCAAAGTGGGAGAACCATCGTGGAACTCTTTTGGGCTTGGGCTGCTGCTGCTG 3001  
|||||  
2963 TCACGCTGTTTCTAGCTTCCAGTGCAGGCTCCAGCTCCAGCTGCTGCTGCTGCTGCTGCT 3022  
|||||  
3002 TCACGCTGTTTCTAGCTTCCAGTGCAGGCTCCAGCTCCAGCTGCTGCTGCTGCTGCTGCTG 3061  
|||||  
3023 CAAATCTTCTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 3082  
|||||  
3062 CAAATCTTCTGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCTTCCAG 3121  
|||||  
3083 CAGCTGTTAGGAGAACTCAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3142  
|||||  
3122 TCAGCAAGTTTGGAGAACCCACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3181  
|||||  
3143 CTGCTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3193  
|||||  
3182 CTGCTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3241  
|||||  
3194 TGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3253  
|||||  
3242 CGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3301  
|||||  
3254 GGTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3313  
|||||  
3302 GACTGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3361  
|||||  
3314 GCTGTCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3373  
|||||  
3362 GCTGAGTCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3421  
|||||  
3374 CCTAAGCAGACGCTTTCAGACCAATTTTGACTAA 3407  
|||||  
3422 ACTGCTCTGAGCTTTCAGACCAATTTTGACTAA 3455  
|||||

AAA29388  
ID AAA29388 standard; cDNA; 4027 BP.  
XX  
AC AAA29388;  
XX  
DT 12-SEP-2000 (first entry)  
XX  
DE hEST2, a human telomerase catalytic subunit homologue cDNA.  
XX  
XX hEST2; telomerase; catalytic subunit; reverse transcriptase; life-span;  
KW retinoblastoma; p53; tumour suppressor; inhibitor; arteriosclerosis;  
KW proliferation; immortal; tumour therapy; macular degeneration; activator;  
KW INK4; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT CDS 57..3455  
FT /\*tag= a  
FT /product= human\_telomerase\_catalytic\_subunit  
XX  
XX WO200031238-A2.  
XX  
XX 02-JUN-2000.  
XX  
XX 24-NOV-1999; 99WO-US27907.  
XX  
XX 25-NOV-1998; 98US-0109891.  
XX 17-FEB-1999; 99US-0120549.  
XX  
XX (GENE-). GENETICA INC.  
XX  
XX Hannon GJ, Beach DH;  
XX  
XX WPI: 2000-400055/34.  
XX P-F5DB; AAY96566.  
XX  
XX New method for increasing the proliferative capacity of cell lines  
XX comprises administering agents reversibly activating telomerase  
XX activity and reversibly inactivating Rb/INK4 and/or p53 pathways useful  
XX in treating age related diseases  
XX  
XX Claim 14; Page 112-116; 123pp; English.  
XX  
XX This cDNA, designated hEST2, is a human telomerase catalytic subunit  
XX homologue of yeast ESR2p and Euplotes p123. hEST2 is a member of the  
XX reverse transcriptase family of enzymes.  
XX The invention concerns methods and reagents for extending the life-span,  
XX e.g. the number of mitotic divisions, of a cell. The method relies on  
XX activation of a telomerase activity and inhibition of one or both of a  
XX retinoblastoma (Rb)/INK4 pathway or a p53 pathway. Phosphorylation of Rb  
XX by cyclin-dependent kinases, cdk4 and cdk6, releases the cells into the  
XX division cycle. Binding of INK4 family members, e.g. the tumour  
XX suppressor p16INK4a, inhibits kinase activity and reversibly inactivate an  
XX arrest. Rb inactivators can selectively and reversibly inactivate an  
XX Rb/INK4 pathway, especially an Rb/p16INK4a pathway. The oncoprotein MDM2  
XX is a cellular inhibitor of Rb/E2f function and the p53 tumour suppressor  
XX and can also be used in the methods. Other molecules which can be used  
XX include cdk4 or cdk6 mutants. In particular, a cdk4 mutant is one which  
XX differs from at one or more of residues K22, R24, H95 and/or D97.  
XX Additional constructs include a papilloma virus E7 protein, or other  
XX viral oncoprotein which bypasses Rb and/or p53. Antisense constructs of  
XX the Rb and p16INK4 genes may also be used. The methods are useful for  
XX increasing the proliferative capacity of cells. The cells are  
XX subsequently of use in pharmaceutical and cosmetic preparations used to  
XX treat conditions related to (premature) ageing, e.g. macular degeneration  
XX and arteriosclerosis. The cells can also be used to replace tumour cell  
XX lines in vitro and for studies on biochemical and physiological aspects  
XX of growth and differentiation. Long lived (immortal) cells could also be  
XX of use in the production of normal or genetically engineered  
XX biotechnology products.  
XX  
XX Sequence 4027 BP; 674 A; 1361 C; 1277 G; 715 T; 0 other;



ery Match 45.8%; Score 1600.8; DB 21; Length 4027;  
 et Local Similarity 68.9%; Pred. No. 0;  
 tches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGAGCAACAATGACCCCGCTCTCGTTGCCCGCGGTGCGTCTCTGCTGG 82  
 41 CCGGCGCAACCCCGCGATGCGCGCTCCCGCTGCGAGCGTCTGCTCTGCTGG 100  
 83 CAGCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGGGCGCTGGGGCCGAGGG 142  
 101 CAGCGCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGGGCGCTGGGGCCGAGGG 160  
 143 CAGGCGGCTTGTGCAACCCCGGAGCCCGGAGATCTACCGCACTTTGGTTGCCCAATGCCCT 202  
 161 CTGGCGGCTGTGCGAGCGCGGAGCCCGGCGCTTTCCGCGCGCTGTGCGCTGCT 220  
 203 AGTGTGCAATGACTGGGCTCAGAGCTCCACCTGCGGACCTTTCTTCCATCAGAGTGT 262  
 221 GGTGTGCGTGGCGGAGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 280  
 263 ATCCCTGAAAGCTGTGGCGAGGTGTGCGAGACTCTGCGAGCGCAACGAGAGAA 322  
 281 CTGCTGTGAGAGCTGTGGCGCGAGTGTGCGAGCGCGCGCGCGCGCGCGCGCG 340  
 323 CGTGTGCGCTTTGCGCTTTGAGCTGTGCTTAACGAGCGCAGAGGCGCGCTCCCATGCCCT 382  
 341 CGTGTGCGCTTTGCGCTTTGCGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCT 400  
 383 CACTAGTAGCGGTGCTAGTCTGCGCAACCTGTTATGAGACCTGCTGCTGAGTGG 442  
 401 CACCAACGCGTGGCGAGCTACTCTGCGCAACCTGCGAGCGCGCGCGCGCGCGCG 460  
 443 TGCATGAGTGTACTGTTGAGCGAGTGGCGAGCGAGCTGTGCTGTACTGCTGGCACA 502  
 461 GCGTGGGCGCTGCTGTTGCGCGCGTGGCGAGCGAGCTGTGCTGCTGCTGCGAG 520  
 503 CTGTGCTCTTTATCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 562  
 521 CTGCGCGCTCTTTGCTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 580  
 563 GTACCAAAATTTGTGCGCACAGGATATCTGCGCGCTGCTGCTGCTGCTGCTGCTGCT 622  
 581 GTACAGCTGCGCGCTGCGCACTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 640  
 623 CCGACCGTGGCGAGAAATTTCACTAACCTTAGGTGTTTCAACAGATCAAGAGAGTAG 682  
 641 GCGTCTGGGATGC-----GAAACGCGCGCTGGAACCATAG 673  
 683 TCGCAGGAGCACCGAAACCCCTGGCGCTTGCCTATCTGAGGTACAAAGAGGATCTGAG 742  
 674 CGTAGGAGGCGCGGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 733  
 743 TCTCACCAGTACAAGTGTGCTTACGCTTAAGAGCGCAGATGCTATCTGCTGCCGAGAT 802  
 734 CAGTGTCCAGCGAAGTCTGCGCTTGCCTCAAGAGCGCGCGCGCGCGCGCGCGCTGAGCC 793  
 803 GGAGGAGGAGCCACAGCGAGGTGCTACCAACCCCATCAGGCAATCATGGTGGCCAG 862  
 794 GGAGCGAGCGCGCTTGGGAGGAGGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 853  
 863 TCGTGTGCGCTCCCGAGGAGTGCCTACT-----GCAGAGAAAGATTGTCTTTCTAA 913  
 854 TGACCGTGTGCTGTGCTGTGCTACCTTGCAGACCGCGCGCGAGAACCCCTCTTTGGA 913  
 914 AGGAAGGTGCTGACCTGAGTCTCTC---TGGGTGCTGTGCTGTGTAACACAGGCCAG 970  
 914 GGTGTGCTCTCTGCGACGCGCGCACTCCACCCATCCGTCGCGCGCGCGCGCGCGCG 973  
 971 CTCACAT---CTCTGCTGTACCAACCGCGCAAAATGCTTTTTCAGTCTCAGGCGATTAT 1027  
 974 CCCCCATCCATCGCGCGCACAGCTCTCCCTGGGACAGCGCTTGTCCCCCGGTGTACGC 1033

1028 TGAGCAGACATTTCTTTACTCCAGGAGATGCGCAAGAGCGTCTAAACCCCTCATTT 1087  
 1034 CGAGACCAACATTTCTCTACTCTCTCAGCGA---CAAGGAGAGCTGCGGCCCTCTT 1090  
 1088 CTTACTCAGCAACTCCAGCTTAACCTTGACTGGGGCCAGAGACTGGTGGAGATCATCTT 1147  
 1091 CTTACTCAGCTCTCTAGGCCAGCTGACTGGCGCTCGGAGGCTGCTGGAGACCATCTT 1150  
 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTGTCAGGACACACCGCTCTATCGCTG 1207  
 1151 TCTGGGTTCCAGGCCCTTGATGCGAGGATCTCCCGAGCTTGGCCCGCTGCCCGAGCG 1210  
 1208 ATACTGCAGATCGGCCCTCTGTTCCAAAGCTGTGCTGTAACCAATGACAGTGCATA 1267  
 1211 CTACTGGCAATGCGGCCCTGTTCTGAGCTGCTTGGAAACCAACGCGCAGTGGCCCTA 1270  
 1268 TGTGAGACTCTCTCAGTTCATGAGGTTTGGACAGCAACCAACAGGTGACAGATGC 1327  
 1271 CGGGTGTCTCTCAAGACGCACTGCGCGCTGCGAGTGGCTCAACCCAGCAGCGGTGT 1330  
 1328 CT-----TGAACACCAAGCCCC 1342  
 1331 CTGTGCCCGGAGAGCCCCAGGCTCTGTGGCGGCCCGGAGGAGAGACACAGACCC 1390  
 1343 ACCGCACTCATGATTTGCTCGCTGCGACAGCACTGCTGCGAGGTATATGTTTTCT 1402  
 1391 CCGTGGCTGTGCTGCTGCGCGCAGCAGCAGCGCTGCGAGGTGTACGGCTTCT 1450  
 1403 TCGGGCTGTCTCTGCAAGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1462  
 1451 GCGGCGCTGCTGCGCGGCTGTGCGCGCGCTGCGGCTTGGGCTCCAGGCAACAAGAG 1510  
 1463 CCGCTTTTAAAGAACTTAAAGAGTTCATCTGTTGGGAAATACGGCAAGCTATCAT 1522  
 1511 CCGCTTCTCAGGAACAACCAAGAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1570  
 1523 GCAGAACTCATGTGGAAGATGAAGTAGAGGATTCCTCTGCTGCGAGAGAGTCTGAGCAAGCTGCA 1582  
 1571 GCAGAGCTGAGTGGAGATGAGCGTGGCGGCTGCGTGGCTGCGCGAGAGCGCCAG 1630  
 1583 GAAGGACGCTGCTCCCGCTGCGAGCAGCCTGTGAGGAGAGGATCTGCTGCTACGTTCT 1642  
 1631 GGTGGCTGTGTTCCGCGCGCAGAGACCGTCTGCTGAGGAGATCTGCGCAAGTCTCT 1690  
 1643 GTTCTGCTGATGAGACATACGCTGTGAGCTGCTTAGTCTATCTTTTATCATCAGCA 1702  
 1691 GCATGCTGATGAGTGTGACGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1750  
 1703 GAGCACATTCAGAAAGAACAGGCTCTTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCA 1762  
 1751 GACCACTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1810  
 1763 GAGATTTGAGTCAAGGCAACCTTGAGAGTGGGCTACCGGAGCTGTCAAGAGAG 1822  
 1811 AAGCATTTGAATCAGACAGCACTTGAAGGAGTGGAGCTGCGGAGCTGTGCGAAGCAG 1870  
 1823 GGTGAGGCTCACAGGACACCTTGGCTAGCATGCGCATCTGCGAGCTGCGTTCATCCC 1882  
 1871 GGTGAGGAGTCAAGGAGCAGGCGCGCTGCTGAGCTGCTGAGCTGCTGCTGCTGCTGCT 1930  
 1883 CAAGCCCAACGCGCTGCGGCCCATTTGTGAACATGATGATGATGATGATGATGATGATGAT 1942  
 1931 CAAGCTGAGCGGCTGCGGCCGATTTGTAACATGAGTACGCTGCTGGAGAGCCAGAACGTT 1990  
 1943 GGGCAGAGGAGAGCAGGCGCCGATTTCAACAGCTCTCAAGACTCTCTTCAAGATGCT 2002  
 1991 CCGCAGAGAGAGAGGCGCGAGGCTCTACCTGAGGAGTGAAGGCACTGTTTCAAGCTGCT 2050  
 2003 CAATATGAGCGGCAAAACATCTCTTATGAGGCTCTTCTGTATCTGGGTATGAATGA 2062  
 2051 CAATACGAGCGCGCGCGCGCGCTCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2110  
 2063 CATCTACAGGACCTGCGCGGCGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122

```

2111 TATCCAGAGGCTGGCGACCTTCTGCTGCTGGGGCCAGGACCCGCGCCCTGA 2170
2123 GATGTACTTTGTTAAGCAGATGTGACCGGGGCTATGATGCCATCCGCCAGGTAAGCT 2182
2171 GCTGTACTTTGTCAAGGTGGATGTGAGGGGCGGTACGACACCATCCGCCAGGACAGCT 2230
2183 GGTGGAGGTGTGGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242
2231 CAGGGAGGTATCGCCAGGATATCA---AACCCAGAACACGTACTCGGTGCGTGGTA 2287
2243 TGCAGTGGTCCGGAGAGATAGCCAGGCCAAGTCCACAAGTCCCTTTAGGACACAGGTAC 2302
2288 TGCGGTGGTCCAGAGCGCCCATGCGGCACTCGCCAGGCTTCCTAAGAGCCACGCTC 2347
2303 CACCTCTCTGACTCCAGGCATACATGGGCCAGTTCCTTAAGCATCTGCAGGATCAGA 2362
2348 TACCTTTGACAGACCTCCAGCCGTACATGCGACGTTCGTGGCTCACCTGCAG-----GA 2401
2363 TGCAGTGCACCTCAGGAACCTCCGTTGTCTATCAGCAGAGCATCTCTATGAATCAGAGCAG 2422
2402 GACAGCCCGCTCAGGGATGCGGTCTATCAGCAGAGCTCTCTCCCTGATGAGCCAG 2461
2423 CAGAGCGCTTTTGACTTCTTCTGCACTTCTGCGTCAAGTCTCGTAAAGATTGGTA 2482
2462 CAGTGGCCTCTTCGAGCTCTTCTAGCTTCTATGTGCCACACGCGCTCGCATCAGGG 2521
2483 CAGTGTCTATACAGCAGTCCAGGGCATCCCGAGGCTCCAGCCTATCCACCTGCTCTG 2542
2522 CAGTCTCAGTCCAGTCCAGGGATCCCGAGGCTCCATCTCTCAGCTGCTCTG 2581
2543 CAGTCTGTGTTTCGGAGACATGAGAACAAAGTGTGTTGAGGTGTCAGCGGATGGTT 2602
2582 CAGCTGTGCTACGGGCACATGAGAACAAAGTGTGTTGCGGGATTCGCGGGACGGCT 2641
2603 GCTTTACGTTTGTGATGACTTCTGTGTGTGAGCGCTCACTGGACCAAGCAAAAC 2662
2642 GCTCCTCGTGTGGTGAATTTCTGTGTGTGACACCTCACTCACCCTCAGCGGAAAC 2701
2663 CTTCTCAGACACCTGCTCATCGCTTCTGAGTATGGTGCATGATAAACTTCAGAA 2722
2702 CTTCTCAGACACCTGCTCGAGGTGCTCTGATGATGCTCGTGGTGAATTCGGAA 2761
2723 GACAGTGTGAATCTCCTGTGAGCTGTGATCCCTGGGTGAGCTCCATACCAGCT 2782
2762 GACAGTGTGAATCTCCTGTAGAACGAGGCGCTGGGTGGACGGCTTTGTTTCAGAT 2821
2783 GCCTGTCACTGCTGTTTCCCTGGGTGCTGCTGTGCTGGACACTCAGACTTTGGAGGT 2842
2822 GCGGCCCCAGGCTATTCCCTGGTGGCGGCTGCTGCTGGATACCGGACCCCTGAGGT 2881
2843 GTTCTGTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACCTCCAGAG 2902
2882 GCAGAGGACTACTTCCAGTATGCCGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2941
2903 TGTCTTCAAGCTGGGAGACCATGCGAACAGCTCCTGTGCTGCTGCTGGTTCAAGTG 2962
2942 CGGTTTCAAGCTGGGAGAACATGCGTCCAAACCTTTGGGGTCTTGGCGTGAAGTG 3001
2963 TCAGGCTTATTTTACACTTGCAGGTGAACAGCCTCCAGACAGTGTGATCAATATATA 3022
3002 TCAGAGCTGTTTCTGATTTGAGGTGAACAGCCTCCAGAGCGGTGTGCACCAATCTA 3061
3023 CAAGATCTTCTGCTTCCAGGCTACAGTTCCATGATGTGATTCAGTTCCTCTTGA 3082
3062 CAAGATCTTCTGCTGAGGGTACAGTTTACGATGTGTGCTGAGCTCCCATTTCA 3121
3083 CCAGGTTTAGGAGAACCTCAATTTCTTGGGCATCATCTCCAGCCAGCATCTG 3142
3122 TCAGCAAGTTTGGAGAACCCCAATTTTCTCTGCGGTCTCTGTGACACGCGCTCCT 3181
3143 CTGCTATGCTATCTGAGGTCAAGATCAGGAATGACATTAAG-----GCCTC 3193
3182 CTGCTACTCTCATCTCTGAAGCAACGACGAGGATGTCTGCTGGGGCCAGGCGCCGC 3241
3194 TGGCTCTTTTCTCTCTGAGCGGCACATTTGGCTCTGTACAGGCTTCTGCTCAAGCT 3253
3242 CGGCGCTCTGCGCTCCGAGGCGGTGCACTGGCTGTGCAACCAAGCATTTCTGCTCAAGCT 3301
3254 GGTCTCTCATTTCTGTCTATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313
3302 GACTCGACACCGTGTACCTACGTGCGCACCTCTCTGGGTCACTCAGGACAGCCGACGCA 3361
3314 GCTGTGCGGAGTCCAGAGGCGCAATGACCATCTTAAAGCTGCACTGACCCAGC 3373
3362 GCTGAGTGGAGTCTCCGGGAGCGAGCTGACTGCTGAGGCGGAGCCGACCAACCGGC 3421
3374 CTTAAGCAGACACTTTCAGACCATTTTGGACTTA 3407
3422 ACTGCGCTCAGACTTCAAGACCATCTCTGGACTGA 3455

RESULT 14
AAV72117
ID AAV72117 standard; cDNA; 4042 BP.
XX AC AAV72117;
XX DT 24-MAY-1999 (first entry)
XX XX Human catalytic telomerase sub-unit cDNA.
DE DE Human; catalytic telomerase subunit; therapy; diagnosis; hTC; assay;
KW modulator; treatment; inhibit; cellular disorder; death; defect; cancer;
XX ageing; antisense; neoplastic cell; telomerase-related condition;
XX tumour cell; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 53..3461
FT /*tag= a
FT /product= "catalytic telomerase subunit"
XX MO9859040-A2.
XX 30-DEC-1998.
XX 09-JUN-1998; 98WO-EP03468.
XX 14-APR-1998; 98DE-1016496.
XX 20-JUN-1997; 97DE-1026329.
XX 26-MAR-1998; 98DE-1013274.
XX (FARB ) BAYER AG.
XX Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;
XX WPI; 1999-081276/07.
XX P-PSDB; AAW90251.
XX New catalytically active subunit of human telomerase - used in the
XX modulation of telomerase activity, particularly for treating cancer
XX and ageing
XX Claim 4; Fig 1; 76pp; German.
XX This sequence encodes a novel human catalytic telomerase sub-unit
XX (hTC). The encoded protein can be used in screening assays to identify
XX modulators of telomerase and to treat or inhibit cellular disorders,
XX death, defects and/or other pathological processes involving telomerase,
XX particularly cancer and ageing (also suitable for these agents that
XX stimulate, inhibit or mimic the activity of the subunit). Antisense
XX nucleic acids inhibit telomerase action (by binding to specific mRNA),
XX particularly in neoplastic cells and may be expressed in vivo. Antibodies
XX and fragments of the protein, used as probes or primers, are used to
```

diagnose telomerase-related conditions (especially neoplasia) by (i) detecting abnormal levels of the subunit protein in body fluids or tissues or (ii) by measuring the amount of the encoding nucleic acid. Expression of the nucleic acid encoding the subunit mRNA is confined to tumour cells, in contrast to the ubiquitous expression of the telomerase RNA subunit.

Sequence 4042 BP; 584 A; 1364 C; 1277 G; 717 T; 0 other;

Query Match 45.8%; Score 1600.8; DB 20; Length 4042;

Est Local Similarity 68.9%; Pred. No. 0;

atches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

```
23 CCGGGCTTGAGACAAATGACCGCGCTCTCGTCCCGCGGTGCGCTCTCTCTCGG 82
47 CCGGGCAACCCCGGATGCGGGGCTCCCGCTCCCGAGCCGCTGCTCTCTCTG 106
83 CAGCCGATACCGGAGGTGTGGCGGTGGCAACTTTGTGCGCGCTGGGGCCGAGG 142
107 CAGCCACTACCGGAGGTGTGGCGGTGGCAACTTTGTGCGCGCTGGGGCCGAGG 166
143 CAGCGGCTTGTGCAACCCCGGACCCGAGATCTACCGCACTTTGGTGGCCCAATGCT 202
167 CTGGCGCTGTGTGAGCGCGGACCCGCGGCTTTCCGCGCTGTGTGGCCCAATGCT 226
203 AGTGTGATGATCTAGCGGCTCACAGCTCTCCACTGCGCACTTTCTTCCACAGGTCT 262
227 GGTGTGCTGCTTGGGACGACGCGCGCGCCCGCGCGCTCTCTTCCGCGAGGTCT 286
263 ATCCCTAAAGAGCTGTGTGCGAGGTGTGTGAGAGCTCTGCGAGCGCAAGAGAA 322
287 CTGCTTGAAGAGCTGTGTGCGCGGAGGTGTGTGAGAGCTGTGTGAGCGCGGCGA 346
323 CGTCTGCGCTTTTGGCTTTGAGCTGTCTTAAACGAGCGCAGAGCGCGCTCCCATG 382
347 CGTCTGCGCTTTGCGCTTTCGCTGTGTGAGCGGCGCCCGCGGGGCGCTT 406
383 CACTAGTAGCGTGTAGTACTTGTGCCCAACAATGTATTATGAGACCTGTGCTGTG 442
407 CACCACAGCGTGTGCGAGTACCTTGTGCCCAACAACGCTGTGACCGCACTGCG 466
443 TGCATGATGCTACTGTGTAGCGAGTGTGCGCGAGCACTGTGTCTACTCTCTGGCA 502
467 GCGTGTGGGCTGTGTGTGCGCGGTGTGCGAGCACTGTGTCTACTCTGTGCGAG 526
503 CTGTGCTTTTATCTTCTGTGTGCGCGCGAGCTGTGCTTACAGGTGTGTGCTTCC 562
527 CTGCGCGCTCTTTGTGTGTGTCTCCAGCTGCGCTTACAGGTGTGTGCGCGCGCT 586
563 GTACCAATTTGTGCCACCGATATCTGGCCCTCTGTGTGCGGTGTGTACAGCGCCAC 622
587 GTACAGCTGTGCGCTGCCACTCAGGCCCGGCCCGCGCACAGCTGTGTGAGCCGAG 646
623 CCGACCGCTGGGAGGAATTTCACTAAGTTTCTTACACAGATCAAGAGCAGTAG 682
647 GCGTCTGGATGC-----GAAAGCGCTGTGAACCATAG 679
683 TCGCCAGGAGACCGAACCCTGGCTTGGCATCTCGAGGTACAAAGAGCATCTGAG 742
680 CGT CAGGAGGCGGGGTCTCCCTTGGGCTTGGCAGCCCGGGGTGTGAGGAGCGGG 739
743 TCTCACCAGTACAGTGTGCTTTCAGTAAAGAGGCGAGATGTATCTGTCTCCGAGAGT 802
740 CAGTGTCCAGCGGAGTCTGCGTTTCCAGAGGCGCCAGCGCTGGCGCTGCCCTGAGCC 799
803 GGAGGAGGAGCCCAAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862
800 GGAGCGAGCGCGCTTGGGAGGAGGTCTTGGGCGCCACCCGGGAGGAGCGCTGTGAG 859
863 TCCTGCTGTGTCCTCCCGAGGTGCTACT-----GCAGAGAAAGATTGTCTTCTAA 913
860 TGACCGGTGTTTCTGTGTGTGTGTCTGCGCACACCGCGCGAAGAGCCACCTCTTTGA 919
```

```
QY 914 AGAAGGTGTCTGACCTGAGTCTCTC---TGGTCTGTGTCTGTAAACACAAGCCAG 970
DB 920 GGTGGCTCTCTTGGACGCGCACTCCACCCATCCGTGGCGCGCAGCACACAGCGGG 979
QY 971 CTCACAT---CTCTGCTGTACACCCCGCAAAATCCCTTTTTCAGTCTAGGCCATTAT 1027
DB 980 CCCCCATCCATCGCGGCCACACGCTCCCTGGGACACGCTTGTCCCGCGGTGTACGC 1039
QY 1028 TGNAGCAGACATTTCTTACTCCAGGGGAGTGGCCAAAGAGCTCTAAACCCCTCAT 1087
DB 1040 CGAGACCAAGCACTTCTCTTACTCTCTAGGGCA---CAAGGAGAGCTGTGGCGCTCT 1096
QY 1088 CTACTACAGCAACCTCCAGCTTAACCTTGACTGGGCGCCAGGAGACTGTGTGAGATCAT 1147
DB 1097 CTTACTCAGCTCTCTGAGGCCCGCAGCTGACTGGGCTCGGAGGCTCGTGGAGACATCT 1156
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTGTGAGGACACACCGTCTATCGCGTCG 1207
DB 1157 TCTGGGTTCAGGCGCTTGGATGCCAGGCACTCCCGCAGGTTCGCCCGCTGTGCCCGAG 1216
QY 1208 ATACTGGCAGATGCGGCGCTCTTCCAAACAGCTGTGTGTGAACCATGCAGAGTGCCAATA 1267
DB 1217 CTACTGGCAATGCGGCGCTCTTCTCGAGCTGCTTGGGAACCAACGCGCAGTGGCCCTA 1276
QY 1268 TGTGAGATCTCTCAGTCACTTGCAGGTTCGAACACAAACCAACAGGTGACAGATGC 1327
DB 1277 CGGGTGTCTCTCAAGACGCACTGCGCGCTGCGACTGCGGTCAACCCAGCAGCGGTGT 1336
QY 1328 CT-----TGAAACACAGCGCC 1342
DB 1337 CTGTGCGCGGAGAGCCCGCGGCTCTGTGGCGCCCCGAGGAGGAGACACAGACC 1396
QY 1343 ACGCACTCATATGATTTGCTCCGCTGCAACAGAGTCCCTTGGCAGGTATATGTTTTCT 1402
DB 1397 CGTCTGCTGTGTGAGCTGCTCCGCGCAGCAGCAGCGCCCTTGGCAGGTGTACGGCTTCT 1456
QY 1403 TCGGCGCTCTCTCGAGGTGTCTGTCTGTAGTCTCTGSGGTACCGGCAACATGAGCG 1462
DB 1457 GCGGCGCTCTGCTGCGCGGCTGTGCTCCCGCGAGCTCTGCGGCTCCAGGCAACAGAAC 1516
QY 1463 CGCTTTCTTTAAGAACTTAAAGAGTTCTATCTGTTGGGAAATACCGCAAGCTATCAT 1522
DB 1517 CGCTTCTCTCAGNAACCAACAGAGTTCTATCTCCCTGGGAGCATGCAAGCTCTGCT 1576
QY 1523 GAGGAACATGATGTGAGAGATGAAGTAGAGTATGCCACTGGCTCCGCGAGCGCGGG 1582
DB 1577 GCAGGAGCTGAGCTGGAAGATGAGCTGCGGAGCTGCGCTTGGCTGTGCGCAGGCGCAG 1636
QY 1583 GAAGGACCGTGTCCCGCTGACAGACACCGTCTGAGGAGAGGATCTCTGGCTACGTTCT 1642
DB 1637 GGTGTGCTGTGTCCGCGCGCAGACACCGTCTGCTGAGGAGATCTCTGGCCAAAGTTCT 1696
QY 1643 GTTCTGGCTGTAGACACATACGTGTGATCAGCTGTAGGTATCTTTTATCATCACA 1702
DB 1697 GCATGGCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1756
QY 1703 GAGCATTCTCAGAGAAACAGGCTCTTCTTCTACCGTAAAGAGTGTGTGGAGCAAGCTCA 1762
DB 1757 GACCAGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTGTGAGCAAGTTGCA 1816
QY 1763 GAGCATTGTAGTCAAGCAACACTTGTAGAGATGTGGGTACGGGAGCTGTCAAGAGGA 1822
DB 1817 AAGCATTGAATCAGACAGCACTTGAAGAGGTGAGCTGTGGGAGAGCTGTGGAAGCAGA 1876
QY 1823 GGTGAGCATCCAGGACACCTGTAGTACCATGCCATCTGCAGACTGCGCTTCATCCC 1882
DB 1877 GGTGAGGAGCATCGGGAAGCAGCGCCCGCTGTGAGTCCAGACTCCCGCTTCATCCC 1936
QY 1883 CAAAGCCAAAGCGCTGCGGCCCATTTGTGAACATAGTATTATAGCATGGGTACAGAGCTTT 1942
DB 1937 CAGCTGTGCGGCTGCGGCCGATTGTGAACATGAGTACTGCTGTGTGGGAGCAGAACGTT 1996
QY 1943 GGGCAGAGAGAGAGCAGGCCCATTTTACCACCGGCTCTCAGAGCTCTCTTCAGCATGCT 2002
```

1997 CCGCAGAGAAAAGAGGGCCGAGCGTCTACCTCGAGGGTGAAGCACTGTTCCAGCGTCT 2056  
2003 CAACCTATGAGCGGACAAAACATCTCTACCTTATGAGGGTCTTCTGCTACTGAGTGAATGA 2062  
2057 CAACCTACGAGCGGGCGGGCCCGGGCTCTCTGGGGCGCTCTGCTGGGCTTGGACGA 2116  
2063 CATCTACAGACCTGCGGGCGGGCTTGTGCTGGTGTGGTGTCTGACACAGACACCCAG 2122  
2117 TATCCACAGGCTTGGCGCACCTTCTGTCTGGTGTGGGGCCCGAGACCCCGCGCTGA 2176  
2123 GATGTACTTTTAAAGCAGATGTGACCGGGGCTTATGATGCCATCCCGGAGGTAAAGCT 2182  
2177 GCTGTACTTTTCAAGTGGATGTGACGGGCGGTACGACACCATCTCCCGGAGGACAGCT 2236  
2183 GGTGGAGGTGTGCGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2237 CAGGGAGGTATCGCCAGCATCATCA---AACCCAGAACACGTACTCGTGGCTGGTA 2293  
2243 TGCAAGTGGTCCGAGAGATAGCAAGGCCAAGTCCCAAGTCCCTTAGGACACAGGTCAAC 2302  
2294 TGCGGTGTTCCAGAGCGGCCCAATGGGCACTCGGCAAGGCTTCAAGAGCCACGCTC 2353  
2303 CACCTCTCTGACCTCCAGCATACATGAGGCCAGTTCCTTAAGCATCTGAGAGTTCAGA 2362  
2354 TACCTTGACAGACCTCCAGCGGTACATGCGAGAGTTCGTGGCTACCTTGCAG-----GA 2407  
2363 TGCAAGTGCATCAGGAACCTCGTGTCTCATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2408 GACAGCGCGCTGAGGGATGCGGTCTCATCGAGCAGAGCTCTCTCCCTGATGAGGCCAG 2467  
2423 CAGAGCGCTTTTGAATCTCTCTGTCACCTTCTGCTGTCACAGTGTGCTAAAGATGGTGA 2482  
2468 CAGTGGCGCTCTTCAGAGCTCTCTACGCTTCTATGTGCCACACGCGGTGCGCATCAGGG 2527  
2483 CAGGTCTATACCACTGACCGGATCCCGAGGCTCCAGGCTTCCAGCTTATCCAGCTGCTGT 2542  
2528 CAAGTCTGATCGTCCAGTCCAGGGATCCCGAGGCTTCCATCTCTCCAGCTGCTGT 2587  
2543 CAGTCTGTGTTTCGAGACATGAGAACAGCTGTTTGTGAGGTGACAGCGGGATGGGTT 2602  
2588 CAGCGTGTCTACGGCGACATGAGAACAGCTGTTTGGGGATTCGGCGGACGGGCT 2647  
2603 GCTTTTACGTTTGTGATGACTTCTGTGTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2648 GCTCCTGCGTTTGGTGGATGATTTCTGTGTGAGACCTTCACTTCACTCCAGCGGAAAC 2707  
2663 CTTTCTCAGCACCTCGTCCATGGCGTTCCTGAGTATGGGTGCATGATAAACTTGCAGAA 2722  
2708 CTTTCTCAGGACCTCGTCCGAGGTGTCCCTGAGTATGCTGCTGGTGAATTCGGGAA 2767  
2723 GACAGTGGTGAATCTTCTGTTGGAGCCTGGTACCTGGGTGGTGGTGCAGTCCATACCAGCT 2782  
2768 GACAGTGGTGAATCTTCTGTTAGAACGAGGCGCTGGGTGGCAGCGCTTGTGTTTCAGAT 2827  
2783 GCTGCTCACTGCTGTTTCCCTGTTGGTGTGCTGCTGAGACCTCAGACTTGTGGAGGT 2842  
2828 GCGGGCCAGCGCTATTCCTGTGTGGGGCTGCTGCTGGATACCCGACCTTGGAGGT 2887  
2843 GTTCTGTGATCTACTAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACTTCCAGAG 2902  
2888 GCAGAGCGATCTACTCCAGTATGCCCGGACCTCCATCAGAGCAGTCTCACTTCAACCG 2947  
2903 TGTCTTCAAGCTGGGAACCATGCGGAACAGCTTCTGTGGTCTTGGGTTGAAGTG 2962  
2948 CCGGTTCAAGCTTGGAGGAAACATGCTGCGAACTCTTTGGGGTCTTGGCGCTGAAGTG 3007  
2963 TCACGCTTATTTCTAGACTTGCAGTGAACAGCTTCCAGACAGTCTGCATCAATATATA 3022  
3008 TCACAGCGCTTCTTGGATTTGAGGTGAACAGCTTCCAGAGCGGTGTGCACCAACATCTA 3067  
3023 CAGAGTCTTCTGCTTACGCGCTTACAGGTTTCCATGATGTGTGATTCAGCTTCCCTTGA 3082

Db 3068 CAGATCTCTCTGCTGAGGGGTACAGGTTTTCAGCATGTGTGTCAGCTCCCATTTCA 3127  
QY 3083 CCAGGCTTTTGAAGAACTCTCACTTCTTTCTGGGCATCATCTCCAGCCAAGCATCTTG 3142  
Db 3128 TCAGCAAGTTTGAAGAACTCCCACTTTTCTGCGGCTCATCTCTGACACGCGCTCCCT 3187  
QY 3143 CTGCTATCTCTCTGAGGTCAGAAATCCAGGAATGACACTAAAG-----GCCTC 3193  
Db 3188 CTGCTACTCTCTCTGAGGTCAGAACTCCAGGATGTCTGCTGGGGCCAAAGCGCGCG 3247  
QY 3194 TGGCTCTCTCTCTCTGAGGCGCACATTTGGCTTCTGTACAGGCTTCTGTCTCAAGCT 3253  
Db 3248 CCGGCTCTGCGCTCCGAGGCGGTGCTGCTGCTGTGTCACCAAGCATTTCTGTCAAGCT 3307  
QY 3254 GCTGCTCATCTCTGTCATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
Db 3308 GACTCGACACCGTGTACCTACGTGCGCATCTCTGGGTCACTCAGNACAGCCAGCGCA 3367  
QY 3314 GCTGTCCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGCTGACGTGACCCAGC 3373  
Db 3368 GCTGAGTCGGAAGCTCCCGGGACGACGCTGACTGCTGAGGCGCGCAGCCAAACCGGC 3427  
QY 3374 CCAAGACACAGCTTTCAGACCATTTTGGACTTA 3407  
Db 3428 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3461

RESULT 15  
ABL53711  
ID ABL53711 standard; cDNA, 4070 BP.  
XX  
AC ABL53711;  
DT  
DT 17-JUN-2002 (first entry)  
XX  
XX Human telomerase catalytic subunit hTERT cDNA.  
DE  
XX hTERT; telomerase; reverse transcriptase; immortalisation; human;  
KW vaccine; enzyme; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WC200216555-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 17-AUG-2001; 2001WO-GB03726.  
XX  
PR 17-AUG-2000; 2000GB-0020246.  
PR 17-AUG-2000; 2000US-225734P.  
XX  
PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.  
XX  
PI Jones CJ, Kipling DG, Wilkinson G, McSharry B, Skinner JW;  
XX WPI; 2002-315462/35.  
XX  
PT Novel hTERT-immortalized cell line (human telomerase reverse  
PT transcriptase) useful for human vaccine production and preparation of  
PT antigen, such as a virus or virus-derived agent -  
XX  
PS Example 1; Fig 1; 64pp; English.  
XX  
CC The present sequence is that of hTERT cDNA in plasmid pGRN121.  
CC hTERT is the catalytic subunit of human telomerase. Claimed  
CC immortalised cell lines for use in vaccine production are adapted  
CC to express hTERT. Suitable cell lines comprise human diploid  
CC fibroblasts, e.g. MRC-5 or WI38 cells, transfected with hTERT cDNA  
CC or infected by a retrovirus carrying hTERT cDNA, and are capable  
CC of supporting antigen production. A method for preparing such cell  
CC lines using recombinant techniques is provided. The cell lines  
CC are also used as a diagnostic test for the presence of a virus,  
CC such as human cytomegalovirus, and to determine the efficacy of





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic - nucleic search, using sw model

on: February 1, 2004, 21:57:16 ; Search time 189 Seconds  
(without alignments)  
8164.420 Million cell updates/sec

le: US-09-042-460-1  
fect score: 3496  
uence: 1 GAATTCGGGTGGAGGCC.....CCGAGTCGGTACCAAGCTT 3496

ring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

ched: 569978 seqs, 220691566 residues

al number of hits satisfying chosen parameters: 1139956

inum DB seq length: 0  
inum DB seq length: 2000000000

t-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : Issued Patents NA.\*  
1: /cgm2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgm2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgm2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgm2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgm2\_6/ptodata/1/ina/PCUTUS\_COMB.seq.\*  
6: /cgm2\_6/ptodata/1/ina/backfileseq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1600.8	45.8	4015	3	US-08-851-843A-224
2	1600.8	45.8	4015	3	Sequence 224, App
3	1600.8	45.8	4015	3	Sequence 1, Appli
4	1600.8	45.8	4015	3	Sequence 224, App
5	1600.8	45.8	4015	4	Sequence 224, App
6	1600.8	45.8	4015	4	Sequence 3, Appli
7	1600.8	45.8	4015	4	Sequence 1, Appli
8	1600.8	45.8	4015	4	Sequence 1, Appli
9	1600.8	45.8	4015	4	Sequence 1, Appli
10	1600.8	45.8	4015	4	Sequence 3, Appli
11	1596	45.7	4037	3	Sequence 173, App
12	1520.2	43.5	4029	3	Sequence 173, App
13	1520.2	43.5	4029	3	Sequence 173, App
14	1520.2	43.5	4029	3	Sequence 173, App
15	1520.2	43.5	4029	3	Sequence 173, App
16	1323.2	37.8	3855	4	Sequence 4, Appli
17	1323.2	37.8	3855	4	Sequence 4, Appli
18	1138.2	32.6	3396	3	Sequence 638, App
19	1116.8	31.9	3451	3	Sequence 721, App
20	1112.4	31.8	3396	3	Sequence 639, App
21	939	26.9	3396	3	Sequence 642, App
22	896.4	25.6	3396	3	Sequence 641, App
23	786.6	22.5	3396	3	Sequence 640, App
24	716	20.5	2176	3	Sequence 3, Appli
25	716	20.5	2176	4	Sequence 3, Appli
26	689.8	19.7	2171	3	Sequence 100, App
27	689.8	19.7	2171	3	Sequence 266, App

Sequence 100, App  
Sequence 100, App  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 30, Appli  
Sequence 62, Appli  
Sequence 8, Appli  
Sequence 62, Appli  
Sequence 62, Appli  
Sequence 31, Appli  
Sequence 9, Appli  
Sequence 213, App  
Sequence 139, App  
Sequence 48, Appli  
Sequence 60, Appli

ALIGNMENTS

RESULT 1  
US-08-851-843A-224  
; Sequence 224, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Hatley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6093809el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:



## SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hprt"

OTHER INFORMATION: /note= "human telomerase reverse

OTHER INFORMATION: transcripase (hprt) catalytic protein

OTHER INFORMATION: component"

08-851-843A-224

Query Match 45.8%; Score 1600.8; DB 3; Length 4015;

est Local Similarity 68.9%; Pred. No. 0;

atches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGGCCTTGAGCAATGACCCGCGCTCCTCGTTGCCCCGCGTGGCTCTCTGTGCG 82  
40 CCCGGCCACCCCGGATGCCGCGCTCCCGCTGCCGAGCGCTGCGCTCCCTGCTGG 99  
83 CAGCCGATACCGGAGGTGTGGCCGCTGGCAACTTTGTGCGCGCGCTGGGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGGCGCTGGCCAGTTGCTGCGCGCGCTGGGCCCGCAGGG 159  
143 CAGCGCGCTTGTCAACCCGGGACCCGAGATCTACCGCACTTTGGTTGCCCAATGCT 202  
160 CTGGCGCTGTGTGAGCGCGGGACCCGCGGCTTCCGCGCGCTGTGTGCCCAATGCT 219  
203 AGTGTGATGCACTGGGGCTCAGAGCTTCAAGCTTCAAGTGTGCTTCAACAGGTGTC 262  
220 GGTGTGCTGCTGGGACGACAGCGCGCGCGCGCGCGCGCTTCCCTTCCGCGAGGTGC 279  
263 ATCCCTGAAGAGCTGTGGCCAGGTTGTGAGAGACTCTGCGAGCGCAACAGAGAAA 322  
280 CTGCTGAAGAGCTGTGTGGCCGAGTGTGAGAGCTGTGAGAGCTGTGAGAGCGCGGAGAA 339  
323 CGTGTGCTTTTGGCTTTGAGTGTCTTAAAGAGGCGAGGCGCGCTTCCATGCGCTT 382  
340 CGTGTGCTTGGCTTGGCTTGGCTTGGAGCGGGCGCGCGCGCGCGCGCGCTT 399  
383 CACTAGTAGCTGTAGTACTTGGCCAGTGGCGAGCACTGCTGTCTTACCTGTGGCACA 442  
400 CACCAACAGCTGCGCACTACTTGCACCAACAGGTGACCGCACTGCGGGGAGCGG 459  
443 TGCATGATGTACTTGTAGCCAGTGGCGAGCACTGCTGTCTTACCTGTGGCACA 502  
460 GCGGTGGGGCTGTGCTGCGCGCGTGGGCGAGAGTGTCTGTTCACTGTGGCAG 519  
503 CTGTGCTCTTTATCTTCTGTGTGCGCGCGAGTGTGCTTACAGGTGTGTGGTCTCCCT 562  
520 CTGCGCGCTCTTGT 579  
563 GTACCAAAATTTGTGCCACAGGATATCTGCGCTCTGTGTGCTGTGTGTGTGTGTGTGT 622  
580 GTACAGCTGTGGCGTGTGCATCTAGGCGCGCGCGCGCGCGCGCGCGCGCGAG 639  
623 CCGACCGGTGGGAGGAATTTCACTTAAGTTTCTTACAGATCAAGAGCAGTAG 682  
640 GCGTGTGGGATG-----GAACGGCGCTTGAACCATAG 672  
683 TCGCCAGAGGACACCGAACCCCTGGCTTCCCATCTCGAGGTACAAAGAGGCATCTGAG 742  
673 CGTCAAGAGGCGCGGGTCCCTCTGGGCTGCCAGCCCGCGGTGCGAGAGGCGCGGGG 732  
743 TCTCAACAGTACAGTGTGCTTCACTAAGAGGCGCAGATGCTATCTTCTCCGAGAGT 802  
733 CAGTGCACCGCAAGTCTGCGGTGCGCAAGAGGCGCGGTGGCGCTGCCCTCTGAGCC 792  
803 GGAGGAGGACCCACAGGAGGTGCTTACCAACCCCATCAGGCAATCATGGGTGCCAG 862

Db 793 GGAGCGGACGCGCGTTGGGCGAGGGTCTCTGGGCCACCCGCGGAGGACCGGTGACCGAG 852  
Qy 863 TCCTGCTGGTCCCGGAGGTGCTACT-----GCAGAGAAAGATTGTCTTTCAA 913  
Db 853 TGACCGTGGTTTCTGTGTGGTGTACCTGCGAGACCGCGCGAAGAGCCACCTCTTTGGA 912  
Qy 914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGGGTGCGGTGTGCTGTAAACACAGCCAG 970  
Db 913 GGGTGGCTCTCTGGCACGCGCCACTCCACCCACTCGTGGCGCCGAGCACACCGCGG 972  
Qy 971 CTCACAT---CTCTGTGTGACACCCCGCAAAATGCTTTTTCAGTCTAGGCCATTTAT 1027  
Db 973 CCCCCATCCATCGCGGCCACACGCTCCCTGGGACACGCTTTGTCCCGCGGTGACG 1032  
Qy 1028 TGAGACGACGATTTCTTTTACTCCAGGGAGATGGCCAGAGGCTTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCACTTCTTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCT 1089  
Qy 1088 CTTACTCAGCAACCTCCAGCTTAATTGACTGGGCGCAGGAGACTGTGTGGAGATCATCT 1147  
Db 1090 CTTACTCAGCTCTCTGAGGCCAGCTGACTGGCGCTCGAGGCTCTGTGGAGCATCT 1149  
Qy 1148 TCTGGCTCAAGGCTTAGACATCAGGACCACTCTGCAAGGACACACCGTCTATCGCGTC 1207  
Db 1150 TCTGGTTTCAGGCGCTTGAATCCAGGAGCTCCCGCAGCTTCCCGCTTCCCGCAGG 1209  
Qy 1208 ATACTGGCAGATGCGGCCCTCTTCCAAAGCTGTGTGAAACCAATGAGAGTGCATA 1267  
Db 1210 CTACTGGCAATGCGGCCCTCTTCTGAGAGCTGTTGGGAAACACAGCGCAGTGCCTA 1269  
Qy 1268 TGTCAAGCTCTCAGTGTACATTTGAGGTTTCGAACAGCAACCAACAGAGTGACAGATGC 1327  
Db 1270 CGGGTGTCTCTCAAGACGCTGCGCGTGGAGCTGCGGTCAACCGAGCGCGTGT 1329  
Qy 1328 CT-----TGAAACACAGGCC 1342  
Db 1330 CTGTGCGGGAGAACCCAGGCTCTGTGCGCGCCCGAGAGGAGGACACAGACCC 1389  
Qy 1343 ACCGACCTCATGAGTTTCTCGGCTGCAAGCACTCCCTGGCAGGTATATGTTTCT 1402  
Db 1390 CCGTCCCTGGTTCAGCTCTCCGACGACAGAGCCCTTGGCAGGTGTACGGTTCGT 1449  
Qy 1403 TCGGGCTGTCTCTCAAGGT 1462  
Db 1450 GCGGCTGTCTGCGCGCTGTGTGCGCGCGCTCTGCGGCTCCAGGCAACAGAACG 1509  
Qy 1463 CCGTCTTTTAAAGATTAAAGATTCTCTGTGTGGGAAATACGGCAAGCTATCACT 1522  
Db 1510 CCGTCTCTCAGGAAACCAAGAGTTTCACTCTCCCTGGGGAAGCATGCCAAGCTCTGCT 1569  
Qy 1523 GCAGGAAGTGTGGAAGATGAAGTAGAGATTGCGCACTGGCTCCGAGCGCCCGG 1582  
Db 1570 GCAAGAGCTGACGTGGAAGATGAGCTGCGGAGCTGCGCTGGCTGCGAGAGGCCAGG 1629  
Qy 1583 GAAGGACCGTGTCCCGCTGCAAGACACCGTCTGAGGAGAGGATCTCTGGCTACGTTCT 1642  
Db 1630 GGTGGTGTGTTCGGCGCGCAGACACCGTCTGCGTGGAGAGATCTCTGGCCAGTTCT 1689  
Qy 1643 GTTGTGCTGATGGAACACATAGCTGTGTAAGCTGTAGTCTAGTCTATTTTATCATCACA 1702  
Db 1690 GCACTGCTGATGAGTGTGTGCTGTGAGTGTGTCTCAGTCTTTCTTTTATGTCCGGA 1749  
Qy 1703 GAGCACAATCCAGAGAACAGGCTCTTTCTTACCGTAAAGTGTGTGGAGCAAGCTCA 1762  
Db 1750 GACCACTTTCAAAAGAACAGGCTCTTTTCTTACCGAAGAGTGTCTGGAGCAAGTGA 1809  
Qy 1763 GAGATTTGGAGTTCAGCAACACTTGTAGAGAGTGGGCTACCGGAGCTGTTCACAGAGA 1822  
Db 1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGACGTGCGGAGCTGTGGAAGACAGA 1869  
Qy 1823 GGTCAAGCATCACAGGACACCTTGGCTAGCCATCTGCAAGACTGCGCTTTCATCCC 1882  
Db 1870 GGTCAAGCATTCGGGAAGGACAGGCCCGCCCTGCTGAGCTCCAGACTCCGCTTTCATCCC 1929

1883 CAAGCCCAACGCGCTCGCGCCCATTTGTGAACATGATTTATAGCATGGGTACCGAGAGCTTT 1942  
1930 CAAGCTCTGACGCGGCTCGCGCCGATTTGTGAACATGGACTACGCTCGTGGGAGCCAGACGTT 1989  
1943 GGGCAGAGGAGGAGGCGGCGGCTTCCAGCCAGCGTCTCAAGACTCTCTTCCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGGCGGCTCTCACTCAGAGGTGAAGCACTGTTTCAAGCTGCT 2049  
2003 CAATATGAGCGGAGCAAAACATCTCTCACTTATGGGGTCTTCTGATCTGGGATGAATGA 2062  
2050 CAATACGAGCGGCGGCGGCGGCTCTCTGCGCGCTCTGCTGGCGCTTGGACGA 2109  
2063 CATCTACAGGACTCGCGGCGCTTCTGCTCGCTGCTGCTCTGAGCCAGACACCCAG 2122  
2110 TATCCACAGGCGCTGGCGACCTTCTGCTGCTGCTGGGCGCCAGGACCGCGCGCTGA 2169  
2123 GATGACTTTTGTAAAGGAGATGTGACCGGGCGCTTATGATGCCATCCCGCAGGATAGCT 2182  
2170 GCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCGCAGGACAGCT 2229  
2183 GTGGAGGTTGTGCAATATGATCAGGCACTCGGAGAGCAGCTACTGTATCCGCCAGTA 2242  
2230 CACGAGGTCATCGCGAGCATCATCA---AACCOCAGAAACAGTACTGCGTGGCTCGGTA 2286  
2243 TGCAGTGTTCGCGAGAGATAGCAAGGCGCAAGTCCCAAGTCTCTTATAGGAGACAGGTCA 2302  
2287 TGCGGTGTTCAGAGAGCGCGCCATGGGCACTCGCAAGGCTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCGAGCCATCATATGAGGCGAGTCTTAAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTATACGACAGTTCGTGGCTCACCTGCGAG-----GA 2400  
2363 TGCAGTGCACCTGAGGAACCTCGTGTGATCGAGCAGAGCATCTCTATGATGAGAGCAG 2422  
2401 GACGAGCGCTGAGGAGTGGCTGCTGATCGAGCAGAGCTCTCTCTGATGAGGCGAG 2460  
2423 CAGCAGCTTTGATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCGAGCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
2483 CAGGTGCTATACGAGTGCAGGCGATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAGTCTTACGCTCAGTGCAGGCGATCCCGAGGCTCCAGGCTATCCACCTGCTGCTG 2580  
2543 CAGTCTGCTTTTCGAGACATGAGGAGCAAGCTGTTGCTGAGTGCAGCGGATGGGTT 2602  
2581 CAGCCTGTCTACGCGGACATGAGAGACAGCTGTTTGGGGGATTCGGCGGACGCGCT 2640  
2603 GCTTTTACGCTTTGTTGATGACTTTCTGTTGTTGAGCGCTCAGCTTGGAGCCAGCAAAAC 2662  
2641 GCTCCTGCGTTTGGTGGATGATTTCTTGTGTTGATGACCTTCACTTCCAGCGGAAAC 2700  
2663 CTTCCTCAGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2722  
2701 CTTCCTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2723 GACAGTGTGAACTTCCCTGCTGAGGCTGTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2782  
2761 GACAGTGTGAACTTCCCTGTTAGAGAGAGGAGGCGCTGCTGCTGCTGCTGCTGCTGCT 2820  
2783 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
2821 GCGGCGCCAGCGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2843 GTTCTGTGACTACTCAGTTATGCCAGACCTCAATTAAGACGAGCTTCACTTCCAGAG 2902  
2881 GCAGAGCGACTACTCAGCTATGCCGAGCTTCCATCAGAGCGAGTCTCACTTCAACCG 2940  
2903 TGTCTTCAAGCTGGAAGACCATCGGAGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTG 2962  
2941 CCGCTTCAAGGCTGGGAGGAAATGCTGCGCAAACTCTTGGGGTCTTGGGCTGGAAGTG 3000

QY 2963 TCACGGTCTATTCTTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGGCTGTTCTTGTGATTTCAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAAGATCTTCTGCTTTCAGGCTTACAGGTTCCATGATGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAGATCTTCTGCTTTCAGGCTTACAGGTTTCAGCATGTGTGCTGAGCTCCCATTTCA 3120  
QY 3083 CCAGGCTTTAGGAAGAACTCCTCATTTCTTTTGGGATCATCTCCAGCAAGCATCTCTG 3142  
Db 3121 TCAGCAAGTTTGGGAAGAACCCCATTTTCTTCCGCTCATCTCTGACACGCGCTCCCT 3180  
QY 3143 CTGCTATGCTATCTGAGGTCAAGATCCAGAAATGACACTAAAG-----GCCTC 3193  
Db 3181 CTGCTATCTCATCTGAAAGCCAGAAACGAGGATGTGCTGGGGGCCAAAGGCGCGC 3240  
QY 3194 TGGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3253  
Db 3241 CCGGCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3300  
QY 3254 GGCTGCTCATTTCTGCTATCTTACAAATGTCTCTTGGGACCTCTGAGGACAGCCCAAACT 3313  
Db 3301 GACTGACACCGTGTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
QY 3314 GCTGTGCGGAGGCTCCAGAGGCGCAATGACCATCTTAAAGCTGCTGAGTACCCAGC 3373  
Db 3361 GCTGAGTGGAGGCTTCCCGGAGCAGCGTGAAGTGGCTGAGGCGCGAGCAACCGCGC 3420  
QY 3374 CTAAGACACAGACTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCTCTCAGACTTCAAGACCATCTCTGGACTGA 3454

## RESULT 2

US-08-974-549A-1  
; Sequence 1, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA: US 08/854,050  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA: US 08/911,312  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: US 08/912,951  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: US 08/915,503  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA: WO PCT/US97/17618  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA: WO PCT/US97/17885  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTRT"  
OTHER INFORMATION: /note= "human telomerase reverse  
OTHER INFORMATION: transcriptase (hTRT) catalytic protein  
OTHER INFORMATION: component"  
80-974-549A-1  
Query Match 45.8%; Score 1600.8; DB 3; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGCGCTTGAGCAATGACCGCGCTCTCGTTGCGCGCGGTGCGCTCTCTGCTGG 82  
40 CCGGCGACCCCGGATGCGCGGCTCCCGCTGCGAGCGGTGCGCTCCCTGCTGG 99  
83 CAGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGGCGCTGGCAACCTTTGTGGCGCGCTGGGCGCCGAGGG 159  
143 CAGCGCGCTTGCAACCGGCGGACCGAAGATCTACCGCACTTTGTTGCCAATGCGCT 202  
160 CTGGCGCGCTGGTGGAGCGCGGCGGACCGCGGCTTTCCGCGCGTGTGGCGCGCTG 219  
203 AGTGTGCATGCACTGGGGCTCAAGCCCTCAACCTTCCCGACCTTTCTTCCACCGAGGTC 262  
220 GGTGTGCGTGGCTGGAGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279  
263 ATCCCTGAAGAGCTGGTGGCGAGGTTGTGAGAGCTCTCGAGCGCGACGAGAGAA 322  
280 CTGCTGAAGAGAGCTGGTGGCGAGGCTGTGAGAGGCTGTGAGAGCGCGCGCGAGAA 339  
323 CGTGTGCGCTTTGGCTTTGAGTGTGTTAAACGAGCGCGAGCGCGCGCTTCCCATGGCTT 382  
340 CGTGTGCGCTTTGGCTTTGAGTGTGTTAAACGAGCGCGAGCGCGCGCTTCCCATGGCTT 399  
383 CACTAGTAGGCTGTAGTACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 442

400 CACCACCGGTGGCGCACTACCTGCTGCCAACACCGGTGACCGACGCACTGCGGGGAGCGG 459  
443 TGCATGAGTACTGTTGAGCGAGTGGGCGACGACTGCTGCTTACCTGCTGGCACA 502  
460 GCGGTGGGGGTGCTGCTGCGCGCGGTGGGCGACGAGTGTGCTGCTGCTGCTGCG 519  
503 CTGTGCTCTTTATCTTCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 562  
520 CTGCGCGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579  
563 GTACCAAAATTTGTGCCACGAGATATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCT 622  
580 GTACCAAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639  
623 CCGACCGGTGGGCGAGAAATTTCACTAACCTTAGTCTTCAACAGATCAAGAGCACTAG 682  
640 GGTCTGGGATGC-----GACGGGCTTGAACCATAG 672  
683 TCGCCAGGAGCACCGAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742  
673 CGTCAAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732  
743 TCTCAGCAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802  
733 CAGTGCAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792  
803 GAGGAGGAGCG 862  
793 GAGCGGAGCG 852  
863 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
914 AGGAAAGGTGTCTGACCTGAGTCTCTC---TGCGTGGTGTGCTGCTGCTGCTGCTG 970  
913 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972  
971 CTCACAT---CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027  
973 CCG 1032  
1028 TGAGACGAGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087  
1033 CGAGACGAGCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089  
1088 CTTACTGAGCACTCTGAGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1147  
1090 CTTACTGAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTCTGAGGACACACCGCTTATCGCG 1207  
1150 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTCTGAGGACACACCGCTTATCGCG 1209  
1208 ATATGCGAGATGCG 1267  
1210 CTTACTGAGCACTCTGAGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269  
1268 TGTGAGCTCTGAGCTCACTGAGCTTTCGAAACGAGCAACCAACAGGTCAGAGATGC 1327  
1270 CCGGCTGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1329  
1328 CT-----TGAAACACGAGCG 1342  
1330 CTGTGCGCGGAGAACCG 1389  
1343 ACCGACATCATGATTTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1402  
1390 CCGTGCCTGCTGAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1449  
1403 TCGGCGCTGCTCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1462  
1450 GCGGCGCTGCTGCG 1509

1463 CGCTTCTTTAAGAACTTAAAGAACTTCACTCTGTTGGGAATACGGCAAGTATCACT 1522  
1510 CGCTTCTCTAGGAACACCAAGAACTTCACTCTGTTGGGAATACGGCAAGTATCACT 1569  
1523 GCAGAACTGATGTGGAAGTGAAGTAGAGATGCACTGCTCGGCAAGAGCCGGG 1582  
1570 GCAGAGCTGAGTGAAGTGAAGTAGAGATGCACTGCTCGGCAAGAGCCGGG 1629  
1583 GAAGACCTGCTCCCGCTCGAGAGCAAGTCTGAGGAGAGATCTGCTGAGTCTCT 1642  
1630 GCTGCTGCTGCTCGGCTCGAGAGCAAGTCTGAGGAGAGATCTGCTGAGTCTCT 1689  
1643 GTTCTGCTGATGAGACACATAGTGTGATGAGTCTGATGAGTCTGATGAGTCT 1702  
1690 GCACTGCTGATGAGTGTGATGAGTGTGATGAGTCTGATGAGTCTGATGAGTCT 1749  
1703 GAGCACTTCCAGAGAGACAGCTCTTCTTCTACCTGTAAGAGTGTGAGAGAGTGA 1762  
1750 GACCACTGTTCAAAAGAACAGGCTCTTCTTCTACCTGTAAGAGTGTGAGAGAGTGA 1809  
1763 GAGCACTGAGTCAAGGCAACACCTTGAAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1822  
1810 AAGCACTGATGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1869  
1823 GGTGAGGATCAGAGAGACCTGCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1882  
1870 GGTGAGGAGATCAGGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1929  
1883 CAAGCCCAACCGCTGCGGCTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 1942  
1930 CAAGCTGAGGCTGCGGCTGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 1989  
1943 GGGCAGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2002  
1990 CCGCAGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2049  
2003 CAAGTATGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2062  
2050 CAAGTATGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2109  
2063 CATCTAGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2122  
2110 TATCCAGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2169  
2123 GATGATCTTTTAAAGAGAGTGTGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2182  
2170 GCTGATCTTTTAAAGAGAGTGTGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2229  
2183 GGTGAGAGTGTGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2242  
2230 CAGGAGAGTGTGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2286  
2243 TGCAGTGTGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2302  
2287 TGCAGTGTGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2346  
2303 CAGCTCTCTGAGCTCCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2362  
2347 TACCTGAGAGAGTGTGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2400  
2363 TGCAGTGTGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2422  
2401 GACGAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2460  
2423 CAGCAGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2482  
2461 CAGTGTGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2520  
2483 CAGTGTGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2542  
2521 CAGTGTGAGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 2580

QY 2543 CAGTGTGAGAGAGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2602  
Db 2581 CAGCTGTGTGTACGGGACATGAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGT 2640  
QY 2603 GCTTTTACGTTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2662  
Db 2641 GCTCTGCGTTTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2700  
QY 2663 CTTCTCAGACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2722  
Db 2701 CTTCTCAGACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2760  
QY 2723 GACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2782  
Db 2761 GACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2820  
QY 2783 GCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2842  
Db 2821 GCGGCTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2880  
QY 2843 GTTCTCAGACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2902  
Db 2881 GACAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2940  
QY 2903 TGTCTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2962  
Db 2941 CCGGCTCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3000  
QY 2963 TCGGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3022  
Db 3001 TCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3060  
QY 3023 CAAGATCTTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3082  
Db 3061 CAAGATCTTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3120  
QY 3083 CAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3142  
Db 3121 TCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3180  
QY 3143 CTGCTATGCTATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3193  
Db 3181 CTGCTATCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3240  
QY 3194 TGCTCTCTCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3253  
Db 3241 CCGGCTCTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3300  
QY 3254 GGTGTGCTATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3313  
Db 3301 GACTCGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3360  
QY 3314 GCTGTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3373  
Db 3361 GCTGTGCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3420  
QY 3374 CTTAAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3407  
Db 3421 ACTGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3454

## RESULT 3

US-08-854-050-224  
; Sequence 224, Application US/08854050  
; Patent No. 6261835  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURES:  
NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTERT"

OTHER INFORMATION: /note= "human telomerase reverse

OTHER INFORMATION: transcriptase (hTERT) catalytic protein

OTHER INFORMATION: component"

:-08-854-050-224

Query Match

Best Local Similarity 45.8%; Score 1600.8; DB 3; Length 4015;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCTTGTGACACATGACCCGCTCTCTGTTGCCCGGGGCGCTCTCTGCTGCG 82

40 CCGGCTTGTGACACATGACCCGCTCTCTGTTGCCCGGGGCGCTCTCTGCTGCG 99

83 CAGCCGATACCGGAGGTGTGGCGCTGTGCAACCTTTGTGGCGGCTGTGGCGCGGAGG 142

100 CAGCCGATACCGGAGGTGTGGCGCTGTGCAACCTTTGTGGCGGCTGTGGCGCGGAGG 159

143 CAGCGGCTTGTGCAACCGGAGGTGTGGCGCTGTGCAACCTTTGTGGCGGCTGTGGCGG 202

160 CTGGCGGCTGTGCAACCGGAGGTGTGGCGCTGTGCAACCTTTGTGGCGGCTGTGGCGG 219

203 AGTGTGATGCTGAGGCTCAGAGCTCCAGCTCGGACCTTTCTTCCACGAGTGTG 262  
220 GGTGTGCTGCTGAGGAGCGACGGCGCGCCCGCCCTCTTCCGCGAGTGTG 279  
263 ATCCCTGAAAGAGCTGTGCTGAGGCTTGTGAGAGACTTCTCGAGCGCAACGAGAGAA 322  
280 CTGCTGAGAGAGCTGTGCTGAGGCTTGTGAGAGCTTGTGAGAGCTTGTGAGAGAGAA 339  
323 CGTGTGCTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 382  
340 CGTGTGCTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 399  
383 CACTAGTAGGCTGCTGAGTACTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGG 442  
400 CACACAGAGCTGCTGAGTACTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGG 459  
443 TGCATGAGTACTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 502  
460 GCGTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 519  
503 CTGTGCTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 562  
520 CTGTGCTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 579  
563 GTACCAATTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 622  
580 GTACCAATTTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 639  
623 CCGACCGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGG 682  
640 GCGTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 672  
683 TCGCAGGAGTACCGGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 742  
673 CCGCAGGAGTACCGGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 732  
743 TCTACAGTACAGTGTGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 802  
733 CAGTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 792  
803 GAGGAGGAGTACCGGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 862  
793 GAGGAGGAGTACCGGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 852  
863 TCTGTGCTGCTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 913  
853 TGACGCTGCTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 912  
914 AGGAAAGGTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 970  
913 GGTGTGCTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 972  
971 CTCCACAT---CTGTGTGTGCTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1027  
973 CCCCCATCCACATGCGGCGCCACCGCTGCTGAGGCTTGTGAGGCTTGTGAGGCTT 1032  
1028 TGAGACGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1087  
1033 CAGAGACGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1089  
1088 CCACTGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1147  
1090 CCACTGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1149  
1148 TGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1207  
1150 TGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1209  
1208 ATACTGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1267  
1210 CTACTGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1269  
1268 TGTGAGTACCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTT 1327

1270	CGGGGTCTCC	TAAGAGCG	ACTGCCCGCTG	CGAGTGGCGGTCA	CCCCAGACGCGGTGT	1329
1328	CT	-----	-----	-----	-----TGAACACAGACCC	1342
1330	CTGTGCCGGG	AGAACCC	CCAGGCGTCTGT	GGCGGCCCCCGAGGAGGAGG	ACACAGACCC	1389
1343	ACCGCACTCAT	GGATTGCTCCG	CTGCGCTG	CGACAGAGTCCCTG	CGCAGGTATATGTTTTCT	1402
1390	CGTCCGCTGG	TGCAGCTG	CTCCGCGACAG	CAGACGCCCTG	CGCAGGTGTACGGCTTCGT	1449
1403	TCGGGCGCTCT	CTCTCAAGGTGG	TGTCTGTAGTCTCTGG	GGTACAGGCA	CAATGAGCG	1462
1450	GCGGCGTCT	CGCGCGCTGT	GCCCGCAGG	CCCTCTGGGCTCCAGG	CAACAGAACG	1509
1463	CGCTTCTTTA	GAACCTTAAGAAG	TTTATCTCTGTTGGG	GAATACGGGAAGCTAT	CACT	1522
1510	CCGCTTCTCT	CAGGAACACCA	AGAAATTCATCTCC	CTGGGGAAGCATG	CCAAAGCTCTCGCT	1569
1523	GCAGGAACGT	ATGTGCGAAGATGA	AAAGTAGAGGATG	CCACTCGCTCCG	CACGACGCCGGG	1582
1570	GCAGGAGCTG	AGCTGCGAAGATAG	CGTGGGACTGG	CGCTTGGCTGCG	AGAGCCGAGG	1629
1583	GAAGGACCGT	GTCCCGCTG	CAGAGCACCGTCTG	AGGAGAGGATCCTGG	CTAGCTTCC	1642
1630	GGTGGCTGT	GTCCGCGCGAG	CAGCACCGCTCTG	CGTGAGGAGATCTCTGG	CAAGTTCCCT	1689
1643	GTCTCGCTGT	ATGGACACATAC	GTGTACAGCTTCTAG	GTCTATTTTACAT	CACAGA	1702
1690	GCACTGGCTG	ATGAAGTGTAC	GTGCTGAGCTGT	CTCAGGTCTTCTTTAT	GTACGGGA	1749
1703	GAGCACATTC	CCAGAAACAGG	CTCTTCTTACC	CTAAGAGTGTGTG	GAGCAAGCTGCA	1762
1750	GACCACTTT	TCAAAAGAACAG	CGCTCTTTTCTACC	CGAAGAGTGTCTG	AGCAAGTTGCA	1809
1763	GAGCATTTG	AGTCAAGCAAC	ACTTGAGAGAGTGG	GGCTACGGAGCTGT	CAAGAGGA	1822
1810	AAGCATTTGA	ATCAGACAGCATTTGA	AGAGGGTGCAGCTCG	GGGAGCTGTG	CGAAGCAGA	1869
1823	GGTCAGGCA	TACACAGGACAC	CTGGCTAGC	CAATGCCCATCTGCAG	ACTGCCCTTCATCCC	1882
1870	GGTCAGGCA	GATCGGGAAG	CGACCGCGCCCTG	CTGACGTCCA	CACTCCGCTTCATCCC	1929
1883	CAAGCCCAAG	CGCTCGGCGCATTTG	GAACATGAGTTATAG	CAATGGGTAC	AGAGCTTT	1942
1930	CAAGCCTGA	CGGGCTCGGCGCGA	TTGTGAACATG	GACTAGCTCGTGG	GAGCCAGAACTT	1989
1943	GGGCAAGAA	GAGCAGGCCAG	CACTTACCCAG	CGCTCAAGACTCTCT	TCACAGTGC	2002
1990	CCGACAGAA	AGAGGGCGAG	CGTCTCACTCG	AGGGTGAAGG	CACTGTTCA	2049
2003	CAACTATAG	AGCGGA	CAAAACATCTCA	ACTTATGGGGTCTT	CTGACTGGGTATGAATGA	2062
2050	CAACTACG	AGCGGGCGGCGCGCG	CGCTCTGCGGCGCTCTG	GTGCGGCGCTCTG	GTGCGGCGCTGGGACGA	2109
2063	CATCTACG	AGCACTCGCGGG	CGCTTTGTGCTGCG	GTGCGGTCTCTG	GCACAGACACCCAG	2122
2110	TATTCAG	AGGGCTGGCGA	CTTCTGCTGCGGTGTG	CGGGCCGAGG	ACCCGCGCGCTGA	2169
2123	GATGTACTTT	GTTAAGGCAGATGT	GACCGGGCGCTATGA	TGCCATCTCC	CCAGGCTAAAGCT	2182
2170	GCTGTACTTT	GTCAAGTGGATGTG	ACGGCGGTACG	ACACCATCCC	CCAGGACAGCT	2229
2183	GGTGGAGGT	TTGGCATATGAT	CAGGCACTCGG	AGACAGCTACTGTAT	TCGCCAGTA	2242
2230	CAGGAGGT	TCATCGCCAGCATCA	CA---AACCC	CCAGAAC	CACGTACTCGTGGCTCGGTA	2286
2243	TGCAGTGG	TCGGAGAGATAG	CCAAGGCCAAG	GTCCAAAGCTCTTT	TAGGACAGAGGTCA	2302
2287	TGCCGTGG	TCAGAGGCGGCCAT	TGGCACGTCCG	AGAGGCTTCA	AGAGCCAGCTTC	2346
2303	CACCTCTCT	GTGACTCGAG	CCATACATGTGG	CGCAGTTCCTTTA	AGCATCTGCAGGATTCAGA	2362

[illegible]

27

RESULT 4  
S-09-430-323-224  
Sequence 224, Application US/09430323  
Patent No. 6309867  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.

TITLE OF INVENTION: No. 6309867el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US 09/430,323

APPLICATION NUMBER: US 09/430,323  
FILING DATE: 29-Oct-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 08-MAY-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"

/note= "human telomerase reverse

transcriptase (hTRT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

JS-09-430-323-224

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGGCTTGAGCACAATACCGGGCTCTGTTGCCCGGGTGGCGCTCTGCTGGC 82

Db 40 CCGGGCACCCTCCGGATGCCGGCGCTCCCGCTGCGAGCGGTGGCTCCCTGCTGGC 99  
QY 83 CAGCCGATACCGGAGGTGTGGCCGCTGGCAACCTTTGTGCGGCGCTGGGGCCCGAGGG 142  
Db 100 CAGCCACTACCGGAGGTGTGGCCGCTGGCCACGTTGTGTGCGGCGCTGGGGCCCGAGGG 159  
QY 143 CAGCGGCTTGTGCAACCCGGGACCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCTT 202  
Db 160 CTGGCGCTGTGCAAGCGCGGGACCCGGGGCTTTCCGGCGCTGGTGGCCAGTGGCT 219  
QY 203 AGTGTGATGCACTGGGGCTCAAGCCTTCACCTGCGGACCTTTTCCTTCCACAGGTGTC 262  
Db 220 GGTGTGGCTGCCCTTGGAGCAGCAGCGCCGCCGCCGCCGCCCTCTCTCCGCAAGTGT 279  
QY 263 ATCCCTGAAGAGCTGTGGCCAGGTTGTGCAGAGACTCTGCAGCGCAACAGAGAGAA 322  
Db 280 CTGCTGAAGAGCTGTGGCCAGGTTGTGCAGAGCTGTGCAGAGCTGTGCAGAGCGGCGAGAA 339  
QY 323 CGTCTGGCTTTTGGCTTTGAGCTGCTTAAACGAGGCGCAGCGGGCTTCCCATGGCTT 382  
Db 340 CGTCTGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTT 399  
QY 383 CACTAGTAGCTGTGGCTAGCTACTTSCCAACACTGTTATTGAGACCTGCGTGTGAGTG 442  
Db 400 CACCACAGCGTGGCAGCTTACCTGCCCAACAGCGGTGACCGACCTGCGGGGAGCGG 459  
QY 443 TGCATGATGCTACTGTGAGCGAGTGGGCGCAGCACTGCTGTGCTTACCTGCTGCGACA 502  
Db 460 GGGTGGGGCTGCTGCTGGCGCGCTGGGCGCAGCAGTGTGTTACCTGCTGCGACG 519  
QY 503 CTGTGCTTTTATCTTGTGTCGCCCGCAGCTGTGCTACCAAGTGTGTGGTCTTCCCT 562  
Db 520 CTGCGCGCTCTTTGTGCTGGCTCCAGCTGCGCTTACCAAGTGTGTGGGCGCGCGCT 579  
QY 563 GTACCAAAATTTGCCACACCGATATCTGGCCCTCTGTGCTCGCTAGTTACAGGCCAC 622  
Db 580 GTACAGCTCGGCGCTCCCACTCAGGCCCGGCCCGCCACACGCTAGTGGAGCCCGAAG 639  
QY 623 CCGACCGCTGGGCGAGAAATTTCACTAACCTTAGTGTCTTACACAGATCAAGAGCAATG 682  
Db 640 CGCTCTGGGATGCTGCTGGCGCTTTCGCTGGCGCGCGCTGGGCGCGCGCTGGGCGCG 672  
QY 683 TCGCCAGGAGACCCGAAACCCCTGGCTTCCATCTCGAGGTACAAAGAGCATCTGAG 742  
Db 673 GTGAGGAGCGCGGGTCTCCCTTGGGCTTCCAGCCCGCGGTGCGAGAGGCGCGGGGG 732  
QY 743 TCTCAGCAGTACAAGTGTGCTTCACTAAGAGGCGCAGATGCTATCTTGTCCCGAGAT 802  
Db 733 CAGTGCACAGCGAAGTCTGCTGGCTTTCGCTGGCGCGCGCTGGGCGCGCGCTGGGCGCG 792  
QY 803 GGAGGAGGACCCCGACAGCGAGTGTACCAACCCCATCAGGCAATCATGGTGCGCAAG 862  
Db 793 GGAGGAGGACCCCGTGGGCGAGGCTCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 852  
QY 863 TCCTGCTGGTCCCGGAGGTGCTACT-----GAGAGAAAGATTTGTCTTAA 913  
Db 853 TGACCGTGGTTTCTGTGTGTGTCACTGCGCAGACCGCGCGAGAGCCACCTCTTGA 912  
QY 914 AGGAAAGTGTCTGACCTGAGTCTCTC---TGCGTGGTGTGCTGTAAACACAGCCAG 970  
Db 913 GGGTGGCTCTCTGGCAGCGCCCACTCCCACTCCCGTGGGCGCGCGCGCGCGCGCGCG 972  
QY 971 CTCACAT---CTCTGCTGTACCAACCCCGCCAAATGCTTTTCAGCTCAGGCCATTTAT 1027  
Db 973 CCCCCATCCACATCGGGCGCACCACTGCTCCCTGGGACACGCTTGTCCCGGTGTACGC 1032  
QY 1028 TGAGACGAGATTTCTCTTCTTCTCAGGAGAGATGGCAAGAGCGCTTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGAGCAGCTGCGGCGCTCTT 1089  
QY 1088 CTTACTAGCAACCTCCAGCTAACTTGTGCTGGGCGCAGGAGACTGTTGAGATCATCTT 1147  
Db 1090 CTTACTAGCTCTCTGAGGCCAGCTGACTGCGCTCGAGGCTCTGAGAGACCATCTT 1149



1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTTCGACGACACACCGCTCTATCGCGTGG 1207  
1150 TCTGGGTTCCAGGCGCTTGATGCGAGGACTCCCGCAGGTTGCGCGCTGCCCGCAGCG 1209  
1208 ATACTGGCAGATCGGCGCTCTGTTTCCAAACAGCTGCTGGTGAACCATGACAGAGTGCCAATA 1267  
1210 CTACTGGAAATGCGGCGCTCTGTTCTGGAGCTGCTTGGAAACCGCGCAGTGCCTCTA 1269  
1268 TGTGAGACTCCTCAGGTCAATTGCAAGTTTCCAAACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGCGTCTCTCAAGACGCACTGCCGCTGCGAGTGGCTCACCACAGAGCGGGTGT 1329  
1328 CT-----TGAAACACGAGGCC 1342  
1330 CTGTGCGCGGAGAGGCCCGCAGGGCTCTGTGGCGGCCCGGAGGAGGACAGAGCCC 1389  
1343 ACCGCACTCATGGAATTTGCTCCGCTGACAGCAGTCCCTGCGAGGTATATGTTTTCT 1402  
1390 CCGTGGCTGTGTCAGCTGTCTCCGACAGCAGCAGCCCTGGCAGGTGTACGGTTGCT 1449  
1403 TCGGCGCTGTCTGCAAGGTGTGTCTGTAGTCTCTGGGGTACGAGGCAAAATGAGGG 1462  
1450 GCGGCGTGTCTGCGCGCTGTGTGCGCCAGGCGCTCTGGGGCTCCAGGCAAAAGGAAAG 1509  
1463 CCGCTCTTTAAAGAACTTAAAGAACTTCACTCTCGTTGGGGAATAGCGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACACCAAGAACTTCACTCTCCCTGGGGAAGCATGCCAAGCTCTCGT 1569  
1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATTTGCACTGGCTCCGAGAGCGCGGG 1582  
1570 GCAGGAGTGAAGTGAAGATGAGCGTGGCGGACTCGCTTGGCTGCGAGAGGCCAGG 1629  
1583 GAAGACCGTGTCCCGCTGCGAGAGCACTGTGAGGAGAGGATCCCTGGCTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGCGCGAGAGCACTGTGCGTGGAGAGATCTGGGCCAAGTTCT 1689  
1643 GTTCTGCTGATGACACATCGTGGTACAGTGTGTTAGTCACTTTTATCATCAGAGA 1702  
1690 GCACTGGCTGATGATGTGATCGTCTGCGAGTGTCTCAGTCTTTCTTTTATGTCAGGA 1749  
1703 GAGCACATCCAGAGAACAGGCTCTTCTTCTACCTGATAGAGTGTGGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTCAAGCAACACCTTGAGAGATGCGGCTACGGGAGCTGTCAAGAGAGA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTGGAAGCAGA 1869  
1823 GGTGAGGATCACCAGACACCTGGCTAGCCATGCCATGCCATCTGCAGACTGCGTTTCAATCCC 1882  
1870 GGTGAGGAGCATCGGAAGCCAGGCGCGCCCTGCTGACGTCAGACTCCGTTTCAATCCC 1929  
1883 CAAGCCCAAACGCGCTGCGGCGCAATTTGAAACATGAGTTATAGCATGGGTACCAGAGCTTT 1942  
1930 CAAGCTGACGCGGTGCGGCGGATTTGTAACATGAGTACGTCGTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGAGAGCAGCGCCAGCAATTTCAACCAAGCTCTCAGACTCTCTTCAGCATGCT 2002  
1990 CCGCAGAGAAAAGAGGCGCGAGGCTCTCACTCGAGGGTGAAGGCACTGTTTCAGGGTGT 2049  
2003 CAATATGAGCGGACAAAATCCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGGCGCGGCGCCCGGCTCTCTGGGGCTCTGTGCTGGGCTGTGACGA 2109  
2063 CATCTACAGGACTGGCGGCGCTTTGTGCGGTGTGCGTGTCTGGACCAAGACCCAG 2122  
2110 TATCCAGAGGCGCTGGCGCACTTCGTGTGTGCGTGTGCGGCGCCAGAACCCCGCGCTGA 2169  
2123 GATCTACTTTTAAAGCAATGTGACCGGGGCTATGATGCAATCCCGAGGATAGCT 2182  
2170 GCTGTACTTTTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCCAGGAGCGCT 2229

2183 GGTGAGGTTGTTGCCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCCGCCAGTA 2242  
2230 CACGGAGTCAATCGCCAGCATCATCA---AACCACAGAAACGCTACTGCGTGTGCGTA 2286  
2243 TGCAGTGGTCCCGAGAGATAGCAAGGCAAGTCCACAAAGTCCCTTTAGGAGACAGGTCCAC 2302  
2287 TGCCTGTGTCCAGAGGCGCGCCCTTGGGCACGTCGCAAGGCCCTTCAAGAGCCACGTCTC 2346  
2303 CACCTCTCTGACCTCCAGGCATACATGGGCGAGTTCCTTAAGCATCTGCAGGATTCAGA 2382  
2347 TACCTTTGACAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCACGTGAGGAACTCCGTTGTCTCATGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCACCCGCTGAGGGATGCCGTGTCATCGAGCAGAGCTCCTCCCTGANTGAGGCCAG 2460  
2423 CAGCAGCTGTTTGTACTTCTTCTGCACTTCTCTGGTCAAGTGTGCTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTTTCGACGCTTCTCTACGCTTCATGTGCCACACGCGCTGCGCATCAGGGG 2520  
2483 CAGGTCCTATACGCAAGTGCAGGCAATCCCGAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAACAAGCTGTTGCTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGTCACGCGACATGAGAAACAAGCTGTTTCGGGGATTCGCGGGGACGCGCT 2640  
2603 GCTTTTACGTTTGTGTGACTTCTGTTGGTACGCTCCTGAGCAAGCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCACCCACGCGAAAC 2700  
2663 CTTCTCTCAGCACCTTGGTCCATGGGCTTCTGAGTATGGGTGCATGATAAACTTGCAGAA 2722  
2701 CTTCTCTCAGGACCTTGTCTCGAGGTGTCCTGAGTATGCTGCGTGGTGAACCTTGGCGAA 2760  
2723 GACAGTGTGAACCTTCCCTGTGAGCCTGTTACCTGGTGGTGGAGCTCCATACCAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTAGAAAGACGAGGCCCTGGGTGGCAGCGCTTTTGTTCAGAT 2820  
2783 GCCTGTCTACTGCTGTTTCCCTGCTGTGGCTGTGCTGTGACACTCAGACTTTTGGAGGT 2842  
2821 GCGGCGCACGCGCTATTCCTCTGTGCGGCTGTGCTGTGATACCGGACCCCTGGAGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAAGACGAGCCTCACCTTCCAGAG 2902  
2881 GCAGAGCACTACTCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAGAACCATGCGGAAACAAGCTCCTGCGGTCTTTCGGTGTGAAGTG 2962  
2941 CGGCTTCAAAGCTGGGAGGAACATGCGTCGCAAACTCTTTCGGGCTTTCGGCTGAAGTG 3000  
2963 TCACGGTCTATTCTTAGACTTCAGGTGAACAGCTCCAGACAGTGTGCATCAATATATA 3022  
3001 TCACAGCTGTTTCTGATTTGAGGTGAACAGCTCCAGACCGTGTGCACCAACATCTA 3060  
3023 CAGATCTTCTGCTTCAGGCTTACAGTTCATGTCATGTCATGTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTCTCTGCTGACGCGTACAGGTTTCAACCATGTCGTGTGAGCTCCCATTTCA 3120  
3083 CAGCGCTGTTAGAAGAACCTCACATTTCTTCTGGGCATCATCTCCAGCCCAAGCATCTGT 3142  
3121 TCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGCTCATCTCTGACACGGCTCTCCCT 3180  
3143 CTGCTATGCTATCTGAGGTCAAGATCCAGGATGACATAAG-----GCCTC 3193  
3181 CTGCTACTCATCTCTGAAAGCAAGACGAGGATGTGCTGGGGGCAAGGGCGCGCG 3240  
3194 TGGCTCTTCTCTGAGAGCGCAGTATGGCTCTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTTCTGCTCTCGAGCGCTGAGTGTGCTGTGCCACCAAGCATTCCTGCTCAAGCT 3300  
3254 GGCTGTCTATCTGTCTCATCTCAAAATGTCTCTCTGGGAGCCTCTGAGGACAGGCCCAAACT 3313

3301 GACTCGACACCGTGTACCTACCTGCGACCTCTCTGGGGTCACTCAGGACGCCACAGCGCA 3360  
3314 GCTGTGCGGAAGCTCCAGAGGCGCAATGACCATCTTTAAAGCTTGCACTGACCCAGC 3373  
3361 GCTGAGTCGGAAGCTCCCGGGAGCAGCTGACTGCTGGAGCGCGCAGCCAAACCCGCG 3420  
3374 CTTAAGCAGACACTTTTCAAGACCACTTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAAGACCACTCTCTGGACTGA 3454

## ULT 5

09-572-423B-3

sequence 3, Application US/09572423B

atent No. 6331399

## GENERAL INFORMATION:

APPLICANT: Brett P. Monia

APPLICANT: William A. Gaarde

APPLICANT: Edward Wancevicz

TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION

FILE REFERENCE: ISPH-0462

CURRENT APPLICATION NUMBER: US/09/572,423B

CURRENT FILING DATE: 2000-05-16

NUMBER OF SEQ ID NOS: 29

SEQ ID NO 3

LENGTH: 4015

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (56)...(3454)

09-572-423B-3

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGCGCCTTGAGCAATATGACCCGCGCTCTCTGTTGCGCGCGGTGGCTCTCTGCTGCG 82  
40 CCGCGCAACCCCGGATGCGCGCGCTCCCGCTGCGGAGCGGTGCGCTGCGTGGCG 99  
83 CAGCGATACCGGAGGTGTGGCGCTGGCACTTTGTGGCGGCTGGGCGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGGCGCTGGCACTTTGTGGCGGCTGGGCGCCGAGGG 159  
143 CAGCGCGCTTGTGCAACCCGCGGACCCGAGATCTACCGCACTTTGTTGCCCAATGCCCT 202  
160 CTGGCGGCTGTGTGACGCGGCGGACCCGCGGCTTTCGCGGCGTGTGGCGCGCT 219  
203 AGTGTGATGTCTGGCGCTTACAGCTTCCACTTCTGCGGACCTTTCTTCCACAGAGTGT 262  
220 GGTGTGCGTGTGCGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279  
263 ATCCCTGAAAGAGCTGTGGCGGCGGCTTGTGACAGACTGTGCGAGCGCAACAGAGAAA 322  
280 CTGGCTTGAAGAGCTGTGTGGCGGCGGAGTGTGCGAGGCTGTGCGAGCGCGCGCGAGAA 339  
323 CBTGTGCTTTTGTGCTTGTAGCTCTTAAAGAGCCAGAGCGCGGCTTCCATGTGCTTT 382  
340 CBTGTGCTTTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399  
383 CACTAGTACGCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 442  
400 CACCCACGCTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 459  
443 TGCATGATGTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 502  
460 GCGGTGCGGCTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 519  
503 CTGTGCTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 562  
520 CTGCGCGCTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 579

QY 563 GTACCAATTTGTGCAACCAAGGATATCTGGCGCTCTGTGTCGCTAGTTACAGGCCAC 622  
DB 580 GTACCAAGCTCGCGCTGTGCACACTCAGGCCGCGCGCGCGCGCACACGCTAGTGACCCGAG 639  
QY 623 CCACCCGCTGGGCGAGAAATTTCACTAACTTAGTTCTTACAACAGATCAAGAGCAGTAG 682  
DB 640 GCCTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672  
QY 683 TCGCCAGGAGACACCGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 742  
DB 673 CGTCAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732  
QY 743 TCTCAGCAGTACAAAGTGTGCTTCAAGTAAAGAGCCAGATGCTATCTCTCTCTCTCTAA 802  
DB 733 CAGTCCAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
QY 803 GGAGGAGGAGACCCCAAGCAGGTGTCTACCAACCCCTCAGGCAATCATGGGTGCCAAG 862  
DB 793 GGAGCGGAGCGCGCTTGGGCGAGGCTCTTGGGCGCGCGCGCGCGCGCGCGCGCGCG 852  
QY 863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
DB 853 TGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
QY 914 AGGAAAGGTGTCTGACCTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTAA 970  
DB 913 GGGTGGCTCTCTGCGCAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 972  
QY 971 CTCCAGTCT 1027  
DB 973 CCCCCCATCCACATCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1032  
QY 1028 TGAGACGAGACATTTCT 1087  
DB 1033 CGAGACGAGCACTTCT 1089  
QY 1088 CTTACTGAGCAACTCTGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1147  
DB 1090 CTTACTGAGCTCTCTGAGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1149  
QY 1148 TCTGGGTCTAAGCGCTTAGGACATCAGGACATCTCTCTCTCTCTCTCTCTCTCTCTCT 1207  
DB 1150 TCTGGGTCTAAGCGCTTAGGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1209  
QY 1208 ATACTGCAAGTGGCGCGCT 1267  
DB 1210 CTTACTGCAAAATGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1269  
QY 1268 TGTGAGACTCTCTGAGGTCTCATGTCAGGTTTCTGAAACAGCAAAACCAAGGTGACAGATGC 1327  
DB 1270 CGGGGTCTCTCTAAGAGGCTCTGCGCGCTCTGAGGTGCGGTCTACCCGAGCGCGGTGT 1329  
QY 1328 CT-----TGAAACACGAGCC 1342  
DB 1330 CTGTGCGCGGAGAACCCCGAGGCTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1389  
QY 1343 ACCGCACTCATGATTTGCTCGCGCTGCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1402  
DB 1390 CGGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1449  
QY 1403 TCGGCGCTGTCTCTGCAAGGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1462  
DB 1450 CGGGGCTGCTGCTGCGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1509  
QY 1463 CGGCTTCTTTAAGTCTTAAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1522  
DB 1510 CGGCTTCTCTGAGGACCAAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1569  
QY 1523 CGAGAACTGATGTGGAAGATGAAAGTAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1582  
DB 1570 CGAGAGCTGACGTGGAAGATGAGCGTGGCGGACTGGCTGGCTGGCTGGCTGGCTGGCT 1629

[illegible]

RESULT 6  
US-09-128-354-1  
Sequence 1, Application US/09128354  
Patent No. 6337200  
GENERAL INFORMATION:  
APPLICANT: Morin, Gregg B.  
APPLICANT: Garon, Corrozzation  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants  
FILE REFERENCE: 015389-003310US  
CURRENT APPLICATION NUMBER: US/09/128,354  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 08/851,843  
EARLIER FILING DATE: 1997-05-06  
EARLIER APPLICATION NUMBER: US 08/854,050  
EARLIER FILING DATE: 1997-05-09  
EARLIER APPLICATION NUMBER: US 08/911,312  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/912,951  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/915,503  
EARLIER FILING DATE: 1997-08-14

EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: US 08/974,549  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 08/974,584  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 09/052,864  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
EO ID NO 1  
LENGTH: 4015  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (56)...(3454)  
OTHER INFORMATION: human telomerase reverse transcriptase (hTRT) cDNA  
09-128-354-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGCGCTTGAGCAATGACCGCGCTCCCTGCTGTCGCGCGCGTGGCGCTCTCTCTGCTGG 82  
40 CCGGCGCAACCCCGGATGCGCGCGCTCCCGCTGCGAGCGCTGGCTCCCTCTGCTGG 99  
83 CAGCCGATACCGGAGGTGGCGCTGGCAACCTTTGTGGCGCGCTGGCGCGCGCGAGGG 142  
100 CAGCCACTACCGGAGGTGGCGCTGGCAACCTTTGTGGCGCGCTGGCGCGCGCGAGGG 159  
143 CAGGCGCTTGCAACCGCGGACCGGAGATCTACCGCACTTTGGTTGCCCAATGCCCT 202  
160 CTGGCGCTTGTTGAGCGCGGGAACCGCGGCTTTTGGCGCGCTGGTGGCGCGCTGCT 219  
203 AGTGTGATGACTCGGCGCTCAGAGCTTCACTGCGCAACCTTTCTTTCACCAAGTGTC 262  
220 GGTGTGCGTCCCTGGGACGACAGCGCGCGCGCGCGCGCGCTCTCTTCGCGAGGTGC 279  
263 ATCCCTGAAGAGCTGGCCAGGTTGTGCAGAGACTCTGCGAGCGCAACAGAGAAA 322  
280 CTGCTGAAGAGCTGGTGGCCGAGTCTGCAGAGGCTGTGCGAGCGCGCGCGAGAA 339  
323 CGTGTGGCTTTTGGCTTTGAGCTCTTAACGAGCGCCAGAGCGCGCGCTCCCATGGCTT 382  
340 CGTGTGGCTTTTGGCTTTGAGCTCTTAACGAGCGCGCGCGCGCGCGCGCGCGCTT 399  
383 CACTAGTAGCGGTAGTACTTCCCAACACTGTATTTAGAACCTTGGTGTGAGTGG 442  
400 CACCACAGCGTGCAGAGTACTTCCCAACACGAGTACCAACCTGCGGGGAGCGG 459  
443 TGCATGATGCTACTTGTGAGCGGAGTGGGAGCGACCTGTGCTTACTGTGGCACA 502  
460 GCGTGGGCGCTGCTGCTGCGCGCGTGGCGAGCGAGTGTGCTTCACTGTGGCAG 519  
503 CTGTGCTCTTTATCTTCTGGTGGCGCGCGAGTGTGCTTACAGAGTGTGGTCTCCCT 562  
520 CTGCGGCTCTTGTGCTGGTGGCTCCAGCTGCGCGCTACAGAGTGTGCGGCGCGCT 579  
563 GTACCAAAATTTGGCCACCAAGATATGCGCGCTCTGTGCTGCTAGTTACAGGCCAC 622  
580 GTACCAAGCTCGGCGTGCACCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639  
623 CCAGCCGCTGGGAGGAAATTTCACTAACTTAGTGTCTTAACAAGATCAAGAGAGTAG 682  
640 GCGTCTGGGATG- 672  
683 TCAGGAGGAGCAGGAGCCCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 742  
673 CGTCAAGGAGCGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 732

QY 743 TCTCAGAGTACAGGTGTCCTTACAGTAAGAGCCAGATGCTATCTCTGTCGAGAGT 802  
DB 733 CAGTGCACCGGAGAGTCTCCCGTTGCCCAAGAGGCCAGCGTGGCGTGCCTTGAGCC 792  
QY 803 GGAGGAGGACCCCAAGAGAGTGTCTACCAACCCCATCAGGCAAAATCATGGTGCCAG 862  
DB 793 GGAGCGGACGCGCTTGGGACAGGCTCTCTGGGCCACCGCGGCGAGGACGCGTGAGCCAG 852  
QY 863 TCTGTCTCGGTCCCGGAGGTGCTTACT-----GCAGAGAAAGATTTTGTCTTCAA 913  
DB 853 TGACCGTGGTTCTGT 912  
QY 914 AGGAAAGGTGTCTGT 970  
DB 913 GGT 972  
QY 971 CTCCAGAT---CTCTGT 1027  
DB 973 CCCCCATCCACATCGCGGCCACCGTCCCTGGGACAGCGCTTGTCCCGCGGTGTACGC 1032  
QY 1028 TGAGACCAGACATTTCTTTTACTTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT 1087  
DB 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGCGCA---CAAGGAGCAGCTGCGGCCCTCT 1089  
QY 1088 CTTACTCAGCAACCTCCAGCCTAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1147  
DB 1090 CTTACTCAGCTCTCTGAGGCCACCGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1149  
QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTGTGAGGACACACCGCTCTATCGCGTCG 1207  
DB 1150 TCTGGGTTCCAGCGCTGTGATGCCAGGACTCCCGCAGTTGCGCGCTGCGCGCGCG 1209  
QY 1208 ATACTGCAATGCGGCGCTCTTTCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1267  
DB 1210 CTACTGCAAAATCGGCGCGCTGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269  
QY 1268 TGTGAGACTCTCAGGTGACATTTGAGGCTTTCGACAGCAAAACCAAGAGTGTGTGTGT 1327  
DB 1270 GGGGCTGCTCTCTAAGAGCACTGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1329  
QY 1328 CT-----TGAAACACAGGCC 1342  
DB 1330 CTGTGCGCGGAGAGGCCCGAGGCTCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1389  
QY 1343 ACCCACTCATGATTTGCTCGCGCTGCACAGCAGTCTCCCTGCGAGGTATATGTTTTCT 1402  
DB 1390 CCGTGGCTGT 1449  
QY 1403 TCGGCGCTGTCTCTGCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1462  
DB 1450 GCGGCGCTGT 1509  
QY 1463 CCGTGTCTTAAAGAACTTAAAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1522  
DB 1510 CCGTGTCTCAGGAAACCAAGAGTTCTCTCCCTGGGAGAGCATGCCAAGCTCTCGCT 1569  
QY 1523 CAGGAACTGATGTGGAAGATGAAAGTAGAGGATGCACTGGCTCCGAGCAGCGCGCGG 1582  
DB 1570 GAGAGGCTGAGTGTGGAAGATGAGCGTGGGAGTGGCTGTGTGTGTGTGTGTGTGT 1629  
QY 1583 GAGAGCAGGTGTCCCGCTGCGAGCAACCGTCTGTGAGGAGAGATCTGTGTGTGTGTGT 1642  
DB 1630 GGT 1689  
QY 1643 GTTCTGGGTGATGAGACATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1702  
DB 1690 GCACTGGGTGATGAGT 1749  
QY 1703 GAGCACAATCCAGAGAACAGGCTCTTTCTTACCGTAAAGTGTGTGTGTGTGTGTGTGT 1762  
DB 1750 GACCACGTTTCAAGAAAGACAGGCTCTTTTCTACCGGAGAGTGTGTGTGTGTGTGTGT 1809

1763 GAGCATTGGAGTCAGGCAACACCTTGGAGAGTGGCGCTACGGAGCTGTCTCAAGAGGA 1822  
1810 AAGCATTGGATATGAGACAGCCTTGAAGAGGGTGCAGCTGCGGAGCTGTTCGAGAGCAGA 1869  
1823 GGTGAGGCATCACAGAGACACCTGGCTAGCCATGCCCATCTCTGCAGACTCGCTTATCCC 1882  
1870 GGTGAGGCAGCATCGGAAGACCAAGCCGCGCTCTGCTGAGCTCCAGACTCGCTTATCCC 1929  
1883 CAAGCCCAAGGGCTGGGCGCCATTTGTGACATGATATAGCATGGTACCAGAGCTTT 1942  
1930 CAAGCTGACGGGTGGGCGCGATTTGGAACATGAGTACGTCTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGCAGGCCAGCATTTTCAACCCAGCGTCTCAAGACTCTCTTCCAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACTCGAGGGTGAAGGCACTGTTCAAGCGTGT 2049  
2003 CAATATGAGCGGCAAAATCATCTCACTTATGGGGTCTTCTGTACTGGGTATGATGA 2062  
2050 CAATAGAGGGCGCGGCGCGCCCTTCTGGGCGCTCTGTGCTGGGCTGGAGGA 2109  
2063 CATCTACAGGACTGGGCGGCTTTGTGCTGCGTGTGCTGCTCTGGACAGACACCCAG 2122  
2110 TATCCAGAGGCTGGCGCACTTCTGCTGCTGCTGCTGGGCGCCAGGACCGCGCTGA 2169  
2123 GATGTACTTTTGAAGGAGATGAGCCGGGCGCTATGATGCCATCCCGCAGGGTAAGCT 2182  
2170 GCTGTACTTTTCAAGGTGATGAGCGGCGGCTAGCAGACCATCCCGCAGGACAGGCT 2229  
2183 GGTGAGGTGTTGCTCCAAATATGATCAGGCACTCGGAGAGCAGTCTGTATCGCCAGTA 2242  
2230 CAGGAGGTATGCCAGCATATCA---AACCCAGAACAGTACTGTGGTGGTGGTA 2286  
2243 TGCAGTGTCCGAGAGATACCAAGGCCAAGTCCAAAGTCTCTTTAGGAGACAGTCA 2302  
2287 TGCCGTGTCCAGAGGCGCGCCATGCGGCAAGTCCGCAAGGCTTCAAGAGGCAAGTCT 2346  
2303 CACCTCTCTGACCTCCAGCATATGAGGCGAGTCTTAAAGCATCTCAGGATTCAGA 2362  
2347 TACCTTGAAGACCTCCAGCGGTATGAGAGTCTGTGGTGCACCTGAG-----GA 2400  
2363 TGCCAGTCACTGAGGAACCTCGTGTGATCGAGAGAGCATCTTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGTGAAGGATGCGTGTGATCGAGAGAGTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGAGGCTGTTGACTTCTTCTGCACTTCTGCGTCAAGTGTGTAAGATGGTGA 2482  
2461 CAGTGGCTCTTTCAGCTCTTCTTACGCTTCTATGTCACCAAGCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGAGTGCCAGGCGATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGTCAGTGCCAGGGGATCCCGAGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTTTTCGAGACATGAGAGACAGTGTGTTGCTGAGGTGACGGGATGGGTT 2602  
2581 CAGCCTGTGCTACGCGCACATGAGAGCAAGTGTGTTGCGGGATTCGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCTTGTGATGATGATTTCTTGTGTTGAGTACCTCACTCACCACCGGAAAC 2700  
2663 CTTCCTCAGCACCTGTGCTCCATGCGTCTGATGATGGTGTGATGAATTAACCTGAGAA 2722  
2701 CTTCCTCAGGACCTGTGCTCCAGGCTGCTCCCTGAGTATGGTGGTGAATTCGGGAA 2760  
2723 GACAGTGTGATCTTCTGAGAGCTGTAACCTGGGTGCTGAGCTCCATACAGCT 2782  
2761 GACAGTGTGATCTTCTGAGAGAGGAGGCTTGGGTGGCAGGCTTGTGTCAGAT 2820  
2783 GCTGCTCACTGCTGTTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCAGGCGCTATTCCTCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGGCTCACCTTCAGAG 2902

2881 GCAGAGGACTACTCCAGTATGCCCGGACCTCCATCAGAGCGAGTCTCACCTTCAACCG 2940  
2903 TGTCTTAAAGCTGGGAGACCATGCGGAAACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGGCTGGGAGGAAACATGCTGCGAAACTCTTTGGGGTCTTTCGCGCTGAAGTG 3000  
2963 TCACGGTCTATTCTAGACTTTCAGGTGAACAGGCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCACAGCTGTTCTGATTTGAGTTCAGGTGACAGCTCCAGAGGCTGTCACCAACATCTA 3060  
3023 CAAGATCTTCTGCTTTCAGGCTTACAGTTCATGTCATGTCATGTCATGTCATGTCATGTCAT 3082  
3061 CAAGATCTTCTGCTTTCAGGCTTACAGTTCATGTCATGTCATGTCATGTCATGTCATGTCAT 3120  
3083 CCAGCGTGTGAGGAGAACCTCACATTTCTTCTGGGATCATCTCCAGCAAGCATCTCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGGCTCATCTCTGACAGGCTCTCCT 3180  
3143 CTGCTATGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3193  
3181 CTGCTACTCATCTCTGAAAGCCAGAAACGAGGATGCTGCTGGGGGCAAGGGCGCGCG 3240  
3194 TGCTCTTCTTCTGAGAGCGCACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3253  
3241 CGGCTCTTCTGCTTTCAGGCGCTGAGTGGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300  
3254 GGCTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3313  
3301 GACTCGACACCGTGTCTACCTAGTGGCTCTCTGCGGCTCACTCAGGACAGCCAGCGCA 3360  
3314 GCTGTGCGGAGCTCCAGAGCGGACATGACCATCTTAAAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 3373  
3361 GCTGAGTGGGAGCTCCCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420  
3374 CCTAAGCACAGACTTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCGCTCAGACTTCAAGACCATCTCTGAGTGA 3454

## RESULT 7

US-09-675-321-1  
; Sequence 1, Application US/09675321  
; Patent No. 6440735  
; GENERAL INFORMATION:  
; APPLICANT: Gaeta, Federico C.A.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune  
; TITLE OF INVENTION: Response to a Telomerase Antigen  
; FILE REFERENCE: 015389-003500PC  
; CURRENT APPLICATION NUMBER: US/09/675,321  
; CURRENT FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/112,006  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898  
; PRIOR FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)  
US-09-675-321-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCGGCGCTTGAGCAGACATACCGCGGCTCTCTGTTGCCCCGGGTGGCTCTCTGCTGCG 82

|||||  
40 CCGCGCCACCCCGCGATGCGCGGCTCCCGCTGCGAGCGTGGCTCCCTGTGTGG 99  
83 CAGCCGATACCGGAGGTGTGGCCGTGCACACCTTTGTGGCGCGCTGGGGCCCGAGG 142  
100 CAGCCACTACCGGAGGTGCTCGCTGGCACGTTCTGTGGCGCTGGGGCCCGAGG 159  
143 CAGCGGCTTTGTGAACCCCGGGACCCGAAGATCTACCGACTTTGTGTGCCCATGCT 202  
160 CTGCGCGCTGTGTGACGCGGGGACCCGCGGCTTTCCGCGCGTGTGTGCCCGCAGTGC 219  
203 AGTGTGATGACACTGGGGCTCAGAGCTCCACTCGCGACCTTTCCCTTCACACAGGTGC 262  
220 GGTGTGTGCCCTTGGACCGACCGCGCGCCCGCGCGCCCTCTCTTCGCCAGGTGTC 279  
263 ATCCCTGAAGAGCTGTGTGCCAGGTTTGTGAGAGACTCTGCGAGCGCAACAGAGAAA 322  
280 CTGCTGAAGAGCTGTGTGCGCGAGTGTGTGAGAGGCTGTGTGAGCGCGCGCGAAGA 339  
323 CGTCTGGCTTTTGGCTTTGAGCTGCTTAAAGAGCGCAGAGCGGGGCTCCCATGCGCTT 382  
340 CGTGTGCGCTTTCGGCTTCGCGTGTGTGACGGGGCGCGCGGGGGCCCCCGAGGCTT 399  
393 CACTAGTAGGTGCGTAGCTACTTGCCCAACTGTATTGAGACCTCGGTGTCAAGTGG 442  
400 CACCACGAGCGTGTGAGCTTACTCTGCCACACGCTGACCGACACTGCGGGGGAGCGG 459  
443 TGCATGATGCTACTGTGTGAGCGAGTGGGGCGAGCACTGTCTGTCTACTGTGTGGACA 502  
460 GCGGTGGGGGCTGTGTGTGGCGCGTGGGGCGACACGTTGTGTGTTCACCTGTGTGCG 519  
503 CTGTGCTCTTTATCTCTGTGTGCCCGCCAGCTGTGCTACCAAGGTGTGGGTCTCCCT 562  
520 CTGCGCGCTTTGTGTGTGTGCTCCAGCTGCGCTTACCAAGTGTGGGGCGCGCT 579  
563 GTACAAATTTGTGCCACCGAGATATGTGGCCCTCTGTGTCCGTGTAGTTCAGGCCAC 622  
580 GTACAGCTCGCGCTGCCACTCAGGCCGGGGCGCGGCCACACAGCTAGTGGACCCGAG 639  
623 CGAGCCGTGGGCGAGGAATTTCACTAACCTTAGTGTCTTACAAAGATCAAGAGCAGTAG 682  
640 GGTGTGGGATGC-----GAA CGGGCTTGAACCATAG 672  
683 TCGCAGGAAGCACCGAAACCCCTTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAG 742  
673 CGT CAGGAGAGCCGCGGTCCCTTGGCTTGCAGCCCGCGGTGCGAGAGGCGCGGG 732  
743 TCTCACAGTACAGTGTGCTTTCAGTAAAGGCCAGATGCTATCTGTCCGAGAGT 802  
733 CAGTGTCCAGCGAAGTCTGCGTGTGCCAAGAGGCCAGCGGTGCGCTGCGCTGAGCC 792  
803 GGAGAGGGACCCACAGGAGGTGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 GAGCGGAGCGCCGTGTGGCAGGGTCTTGGGCCACCGGGCAGGACGCGTGAACCGAG 852  
863 TCCTGCTCGTCCCGCAGAGTGCCTACT-----GCAGAAAGATTGTCTTCTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCACTGTGCAGACCGCGCGAAGAGCACCTCTTTGGA 912  
914 AGGAAAGTGTCTGACCTGAGTCTCTC- --TGGGTGGGTGCTCTGTAAACACAGGCCAG 970  
913 GGGTGGCTCTCTGGCAGCGGCACCTCCACCCATCCGTGGGCCCGCAGACACACGGGG 972  
971 CTCCACAT- --CTGTGTGTACACACCGGCCAAAATGCTTTCAGCTCAGGCCATTTAT 1027  
973 CCCCCATCCATCGCGGCCACCCACGCTCCCTGGGACAGCGCTTGTCCCGCGGTGACGC 1032  
1028 TGAGACCGAGCATTTCTCTTTACTCCAGGGAGATGCGCAGAGCGTCTAAACCCCTCAT 1087  
1033 CAGACACAGACATTTCTCTTACTCTCCTCAGCGCA- --CAGAGGAGAGCTTGGGGCTCTCT 1089  
1088 CCTACTCAGCAACCTTCAGCGCTTAATTGACTGGGGCGAGGACATGGTGGAGATCATCTT 1147

1090	CTTACTCAGCTCTCTGAGGCCCGACCTGACTGGCGCTCGAGGCTCTGTGGAGACCATCTTT	1144
1148	TCTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGACGACACACCGCTCTATCGCGTCG	1207
1150	TCTGGTTCCAGGCCCTGGATGCGAGGACATCCCGCAGGTTGGCCCGCTTGCCCGCAGCG	1209
1208	ATACTGGCAGATGCGGCCCTGTCTCCAAAGCTGCTGTGTAACTATGACCATGACAGAGTGC	1267
1210	CTACTGGCAAAATGGGCCCTGTCTTCTGGAGCTGTCTTGGAAACCAAGCGCAGTGC	1269
1268	TGTCACTCTCTCAGGTCACATGTCAGGTTTCGAAACAGCAAAACCAACAGGTGTACAGATGC	1327
1270	CGGGGTGCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTACCCAGCAGCGGCTGT	1329
1328	CT-----TGAACACCAAGCCC	1342
1330	CTGTGCGGGAGAGCCCGAGGCTCTGTGGCGGCCCGAGGAGGAGGACACAGACCC	1399
1343	ACGCACTCTATGATTTGCTTCGCCCTGACACAGCTCCCTGGCAGGTATATGTTTTCT	1402
1390	CGCTGCCCTGTGTGCTGCCAGACACAGCAGCCCTGGCAGGTGTACGGCTTGT	1449
1403	TGGGCTGTCTCTCAAGTGTGTCTGTACTGTCTGGGTACAAGCACAATGACGC	1452
1450	CGGGCTGCTCTCGCCGCGCTGTGTGCCCGCAGGCTCTGGGCTCAGGCAACAGAACG	1509
1463	CGCTCTTTTAGAATTTAAGAGTTTCATCTGTTGGGGAATACGCAAGCTATCACT	1522
1510	CGCTTCTCTAGAAACCAAGAAGTTTCACTCTCCCTGGGAGAGATGCCAGCTCTCGT	1569
1523	GCAGGAATGATGTGCGAAGATGAAAGTAGAGATTCCCATCGCTCCGACAGACCCGGG	1582
1570	GCAGGAGCTGACGTGGAAGATGAGCTGCGGACTCGCTGGCTGGCAGGAGCCACG	1629
1583	GAAGGACCTGTCCCGCTGCAGAGACCGCTCTGAGGAGAGATCCTGGCTAGCTTCCCT	1642
1630	GGTTGCTGTGTCTCGGCCGACAGACACCGTCTGCTGTAGAGATCTCTGCCAGTTCCT	1689
1643	GTCTCGGCTGATGACATACGCTGTACAGCTCTTAGGTCACTCTTTTATCATCACAGA	1702
1690	GCATCGCTGATGATGTGTACGTCGTGAGCTCTCAGCTCTTCTTTATGTACGGA	1749
1703	GAGCAATTCAGAGAAACAGGCTCTTCTTCTACGTAGAGATGTGTGGAGCAAGCTGCA	1762
1750	GACACGTTTCAAAAGAACGGCTCTTTTCTACCGGAAAGATGTCTGAGCAAGTGCA	1809
1763	GACATTTGAGTTCAGCAACACTTTGAGAGAGTTCGGCTACGGAGCTGTCAAGAGGA	1822
1810	AAGATTGAATTCAGACAGCACTTGAAGAGGTGTGAGCTGCGGAGCTGTGGAAGCAGA	1859
1823	GGTCAGGCATCACAGGACACTGGCTAGCCATGCCCATCTGACAGCTTCGCTTCATCCC	1882
1870	GGTCAGGACGATCGGGAAGCAGGCCCGCCCTGCTGACGCTCCAGACTCCGCTTCATCCC	1929
1883	CAGGCCAACGCGCTCGGCCCATTTGTGAACATGATTAAGCATGGGTACCAAGCTTTT	1942
1930	CAGGCTGACGGGCTCGGGCGATGTGAACATGAGCTACGCTGTGGAGAGCCAGACGTT	1989
1943	GGGCAGAGAAAGCAGGCCACAGCAATTCACCCAGCGTCTCAAGACTCTCTTCAGCATGCT	2002
1990	CCGCAGAAAGAGGCGGAGGCTCTCACTCGAGGTTGAAGCACTGTTCAGCGTCTCT	2049
2003	CAACTATGAGCGGACAAACATCTCACTTATGGGCTCTTCTGTACTGGGTATGAATGA	2062
2050	CAACTCAGAGCGGCGCGCGCCCGGCTCTCTGGCGCTCTGTGCTCTGGGCTCTGACGA	2109
2063	CATCTACAGGACTCGCGGGCTTTGTGCTCGGTGTGCGTGTCTGGAACAGACACCCAG	2122
2110	TATCCACAGGCTTGGGCACTTCTGTGCTGTGGGTGTGGGCCCCAGGACCCCGCGCTGA	2159
2123	GATGTACTTTGTTAAGGAGATGTGAACGGGGCTATGTATGCGCATCTCCCGAGGTTAAGCT	2182
2170	GCTGTACTTTGTCAAGGTGATGTGAACGGGCGGCTACGACCACTATCCCGACAGACGCT	2239

2183 GGTGAGGTTGTTGCCAATATGATGAGCACTCGAGAGACAGTACTGTATCGCCAGTA 2242  
2230 CACGAGGTGATCGCCAGCATCA---AACCCCAAGAACAGTACTGCGTGGT 2286  
2243 TGCAGTGTCCGAGAGATAGCCAAAGCCAAAGTCCCTTTAGGAGACAGGTTCAC 2302  
2287 TGCCTGTGTCAGAGGCGCCCATGGCACGTCCGCAAGGCTTCAAGAGCCACGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCTTAAAGCATCTGAGATTCAGA 2362  
2347 TACCTTGAAGACCTCCAGCCGATACATGCGACAGTTCTGCGCTCACTGCGAG-----GA 2400  
2363 TGCCAGTGCATGAGGAACCTCCGTTGTTCATGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGTGGAGGATCCGTCGTCATGAGCAGAGCTCTCTCCCTGAATGAGCCAG 2460  
2423 CAGCAGCTGTTTACTTCTCTGACCTCTCTGCTGTCAGAGTGTGTAAGATTGGTGA 2482  
2461 CAGTGGCTCTTCGACGCTTCTCTACGCTTCTCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGTGTCTATAGCAGTCCAGGCTCCAGGCTCCAGGCTCCAGCTATCCACCTGCTCTG 2542  
2521 CAGTCTCTAGTCCAGTCCAGGCTCCAGGCTCCAGGCTCCAGTCTCTCTGCTGCTGCTG 2580  
2543 CAGTGTGTTTCCAGACATGAGGAACAAGCTGTTTCTGAGGTCCAGCGGATCGGTT 2602  
2581 CAGCCTGTCTACGCGACATGAGGAACAAGCTGTTTCCGCGGATTCGCGGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGTTGTCAGCGCTCACTGAGCCCAACAAC 2662  
2641 GCTCTGCTTGTGAGGATGATCTTCTGTTGTCAGCTCACTGAGCCCAACAAC 2700  
2663 CTCTCTAGCCTCTGTCATGCGCTCTGAGTATGGTGTGATGATGAATGAGGAA 2722  
2701 CTCTCTAGGACCTCTGTCAGGCTCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTG 2760  
2723 GACAGTGTGATCTCTGTCAGGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCT 2782  
2761 GACAGTGTGATCTCTGTCAGGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCTGTCAGCT 2820  
2783 GCTCTCTCACTGCTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
2821 GCGGCCCCAGGCTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2843 GTTCTGTGATCTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACTTCCAGAG 2902  
2881 GCAGAGCAGCTACTCCAGTATGCCGAGCTCCATCAGAGCAGCTCACTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGAGACCATGCGGAACAAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2962  
2941 CGGCTTCAAGCTGGAGAACATGCTGCGAACAATCTTTGGGCTTTCGCGCTGAAGTG 3000  
2963 TCACGCTTATTTCPAGACTTGCAGGTGAACAGCCTCCAGACAGTGTGCATCAATATATA 3022  
3001 TCACAGCCTGTTTCTGATTTGCAAGTGAACAGCCTCCAGACGCTGTCACCAACATCTA 3060  
3023 CAAGATCTCTCTGCTTCAAGCTTCAAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3082  
3061 CAAGATCTCTCTGCTGAGGCTTCAAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3083 CCAGCCTGTTAGGAAGACCTCACAATCTTTCTGGGATCATCTCCAGCCAAAGCATCTCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATATTTCTCTGCGCTCATCTCTGACACGCGCTCCCT 3180  
3143 CTGCTATCTATCTGAGGTCAAGATCCAGGAATGACACTAAG-----GCTCT 3193  
3181 CTGCTACTCTCTCTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 3240  
3194 TGGCTCTCTTCTCTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 3253  
3241 CGGCTCTCTCTCTGAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCT 3300

QY 3254 GGTGCTCATCTGTCTATCAAAATGTTCTCTGGACCTCTGAGACAGCCCAAAATCT 3313  
DB 3301 GACTCGACACCGTGTACCTTACGTGCTGCTCTCTGGGTCTCTGAGACAGCCAGCA 3360  
QY 3314 GCTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGCTGACAGTGCACCCAGC 3373  
DB 3361 GCTGAGTCCGAAGCTCCGCGGAGCAGCTGCTGCTGCTGAGGCGGCGACCAACCCGCGC 3420  
QY 3374 CCTAAGCACAGACTTTCAGACCATTTTGGACTAA 3407  
DB 3421 ACTGCTCAGACTTCAAGACCATCTCTGACTGA 3454

## RESULT 8

US-09-052-919-1  
; Sequence 1, Application US/09052919  
; Patent No. 644650  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Antisense Compositions for Detecting and  
; TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/052,919  
; FILING DATE: 31-MAR-1998  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/974,549  
; FILING DATE: 19-NOV-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/974,584



FILING DATE: 19-NOV-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Parent, Annette S.

REGISTRATION NUMBER: 42,058

REFERENCE/DOCKET NUMBER: 015389-003600US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "human telomerase reverse

OTHER INFORMATION: transcriptase (hTRT)"

-09-052-919-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 23/9; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCGGGCTTTGAGCACAATGACCGGGCTCTCGTTGCGCGCGCGTCTCTGCTGCG 82

40 CCGGGCAACCCCGCGATGCGGGGCTCCCGCTCCGAGCGGTGCGTCCCTGCTGG 99

83 CAGCGATACCGGAGGTGTGGCGGTGCGAATCTTTGTGTGGCGCGCTGGGCGCGAGG 142

100 CAGCCACTACCGGAGGTGTGCGGTGCGCACTGTTGCGGGCGCTGGGCGCGCGG 159

143 CAGCGGCTGTGCAACCGGGGACCGGAGATCTACCGCACTTTGGTTGCCAATGCT 202

160 CTGGCGGTGTGACGGGGGACCGGCGGCTTTCCGGCGGTGTGGCGCGTGCCT 219

203 AGTGTGATGACATGCGGGTCAAGCTTCCACTGCGGACCTTTCTTCCACAGGTGTC 262

220 GGTGTGCTGCTCGGAGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279

263 ATCCCTGAAGAGCTGTTGCGGAGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322

280 CTGCTGAAGAGCTGTTGCGGAGGTTGTGAGAGGTTGTGCGAGCGCGCGCGAGAA 339

323 CGTGTGCTTTTGGCTTTGAGTGTAAAGAGCGCGAGCGCGCGCTCCCATGGCTT 382

340 CGTGTGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 399

383 CACTAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442

400 CACACAGCGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 459

443 TGCATGATGCTACTGTTGAGCGAGTGGCGGCGAGCGAGCTGCTGCTGCTGCTGCT 502

460 GCGTGGGGCTGCTGCTGCGCGCGTGGCGCGAGCGTGTGCTGCTGCTGCTGCT 519

503 CTGTGCTTTTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562

520 CTGCGCGCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579

563 GTACCAATTTGGCCACACGATATCTGGCGCTCTGTGCTGCTGCTGCTGCTGCT 622

580 GTACAGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 639

623 CCGACCGTGGCGAGGAATTTCTAATCTTAGTGTCTTACCAAGATCAAGAGCAGTAG 682

Db 640 CGGTCTGGATGC-----GACGGGCTGGAACATAG 672

QY 683 TCGCAGGAAGCACCAGAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742

Db 673 CGTCAGGAGGCGCGGGTCCCTTGGCTTGCAGCGCCCGGGTGCAGAGGCGCGGGG 732

QY 743 TCTCAGCAGTACAGTGTGCTTACAGTAAGAGGCGAGATGTAATCTGTCTCCGAGAGT 802

Db 733 CAGTGCCAGCCGAAGTCTGCGTTGCCAAGAGCGCCAGGCTGGCGCTGCCCTGAGCC 792

QY 803 GGAGAGGAGACCCACAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGGTGCAAG 862

Db 793 GGAGCGAGCGCGCTTGGCGAGGGTCTTGGGCGCCACCGGGCAGGACGCTGACCGAG 852

QY 863 TCCTGCTCGTCCCGCGAGGTGCTACT-----GCAGAGAAAGATTGTCTCTTAA 913

Db 853 TGACCGTGGTTTCTGTGTGTGTGTCCTGCGACACCGCGCAAGAGACCACTCTTTGA 912

QY 914 AGGAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGGTGTGTGTAACACAAAGCCAG 970

Db 913 GGTGGGCTCTCTGCGACGCGCACTCCACCCATCGTGGGCGCGCAGCACCGCGG 972

QY 971 CTCACAT---CTGTGTGTACACACCCCGCAAAATGCTTTCAGCTCAGGCCATTAT 1027

Db 973 CCCCCCATCCCATCGCGGCGCACCATCTCCTGGGACACGCTTGTCCCCCGGTGTACGC 1032

QY 1028 TGAGACCAAGCATTTCTTACTCCAGGGAGATGSCCAAGAGCGTCTAAACCCCTCAT 1087

Db 1033 CGAGACCAAGCATTTCTTACTCTCAGGCGA---CAAGGACAGCTGCGGCGCTCT 1089

QY 1088 CTTACTAGCAACTCCAGCTTAACTTGTGCTGAGTGGGCGAGAGACTGGTGGAGATCATCT 1147

Db 1090 CTTACTAGCTCTCTGAGGCGCGAGCTGACTGCGGCTGCGAGGCTGCTGGAGACCATCT 1149

QY 1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTCTGCGAGACACCGCTATCTGGGTG 1207

Db 1150 TCTGGTTTCCAGGCGCTTGTATGCGAGGACTTCCCGAGGTTGCCCGCTGCGCCAGCG 1209

QY 1208 ATACTGGCAGATGCGGCGCTTTCACAGCTGCTGGTGAACCATGACAGGTGCCAATA 1267

Db 1210 CTACTGGCAATGCGGCGCTTCTTCTGAGAGTCTTGGGAGACCGCGCAGTGGCGCTA 1269

QY 1268 TGTGAGATCTCTCAGGTCATCTGAGGTTTTCGAGACAGCAACCAACAGGTGACAGATGC 1327

Db 1270 GCGGCTGCTCTCAAGACGCACTGCGGCTGCGAGTTCGCTCAGCCAGCAGCGGTGT 1329

QY 1328 CT-----TGAAACACAGGCC 1342

Db 1330 CTGTGCGCGGAGAGCCCGAGGGCTCTGTGGGCGCCCGAGGAGGAGGACACAGACCC 1389

QY 1343 ACCGACCTCATGATTTGCTCCGCTGCGACAGAGTCCCTGGCAGGTATATGTTTCT 1402

Db 1390 CCGTCCCTGTGTGAGTGTCTCCGCGAGCAGAGCGCTTGGCGAGGTGTACGGCTTGT 1449

QY 1403 TCGGCGCTGTCTCTGAGAGTGTGTCTGTAGTCTCTGGGGTACCGAGCACAATGAGCG 1462

Db 1450 GCGGCGCTGCTGCGCGGCTGTGTGCGCCAGGCTCTGGGCTCCAGGCAACAGAACG 1509

QY 1463 CCGCTCTTTTAAAGACTTAAAGAGTTCATCTCGTGGGAAATACGGCAAGTATCACT 1522

Db 1510 CCGCTTCTCTAGGAACACCAAGAGTTTCTCTCCCTGGGAAAGCATGCCAAGCTCTGCT 1569

QY 1523 GCAGAACTGATGTGAAGATGAAAGTAGAGATTGCCACTGGCTCCGAGAGCGCGGG 1582

Db 1570 GCGAGGCTGAGTGAAGATGAGGCTGCGGACTGGCTTGGCTGCGCAGGAGCGCGAG 1629

QY 1583 GAAGACCGTGTCCCGCTGCGAGGACCGGTCTCAGGAGAGGATCTCGGTACGTTCTCT 1642

Db 1630 GGTGGCTGTGTTCGCGCGCAGAGCACCGCTCTCGTGAAGAGATCTCGGCAAGTTCTCT 1689

QY 1643 GTTCTGGCTGTGAGACATACGTTGAGTGTACAGTCTCTTAGGTCTTTTACATCAGA 1702

1690 GCACCTGGCTGATGAGTGTGACGTCGTCGAGCTGCTCAGGTCTCTTTCTTTATGTCACGGA 1749  
1703 GAGCACAATTCAGAGAACAAGCTCTTTCTTCTACCGTAAGAGTGTGTGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAAGCTCTTTCTTCTACCGTAAGAGTGTGTGAGCAAGTTCGA 1809  
1763 GAGCATTGGAGTTCAGGACAACACCTTTGAGAGGTGCGGTACGGAGCTGTCAAGAGGA 1822  
1810 AAGCAATTGGAATCAGACAGCACTTGAAGAGGTGACGTCGCGAGCTGTGCGAAGCAGA 1869  
1823 GGTGAGGCATCACCAGGACACCTTGGCTAGCCATGCCCCATCTGACAGACTGCGCTTCATCCC 1882  
1870 GGTGAGGCAGCATCGGGAAGCAGGCGCGCCCTGCTGACGTCACAGACTCGCTTCATCCC 1929  
1893 CAGGCCAACCGCTGCGGCCCATTTGTAACATGAGTTATAGCATGGTACGAGACTTT 1942  
1930 CAAGCCTGACGGGTGCGGCCGATTTGTAACATGAGTTATAGCATGGTACGAGACTTT 1989  
1943 GGCAGAGAAGGAGCAGGCGCCAGCAATTTCAACCAGCGTCTCAAGACTCTTTCAGCATGCT 2002  
1990 CGCAGAGAAAGAGGGCCGAGCGCTCACTCGAGGGTGAAGGCACGTGTCAGCGTGTCT 2049  
2003 CAACTATAGCGGAGCAAAACATCTCACTTATGCGGTCTTCTGACTGGGTATGAAATGA 2062  
2050 CAACTACAGCGGCGCGCGCCCGCCCTCTTGGGCGCTCTGTGTGGGCGCTGAGCA 2109  
2063 CATCTACAGGACCTGGCGGGCTTTGTGTCGCTGCTGCTGCTGAGCAGACACCCAG 2122  
2110 TATCCAGAGGGCTGGCGCACTTCTGTGCTGCTGCTGCTGCGGCCAGGACCGCGCTGA 2169  
2123 GATGATCTTTGTTAAGCAGATGTGAACGGGGCTTATGATGCCATCCCCAGGGTAAGCT 2182  
2170 GCTGTACTTTGTTCAAGTGTGATGTGAACGGGGCTTATGATGCCATCCCCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTCATATATGATCAGCACTCGGAGAGCAGTCTGATCTCGCCAGTA 2242  
2230 CAGGAGGTGATCGCCAGCATATCA--AACCACCAACAGTACTGCTGCTGCTGCTG 2286  
2243 TGCACTGTGCGGAGATAGCCAAAGCCAGTCCCAAGTCTCTTTAGGAGACAGGTCA 2302  
2287 TGCGGTGTCCAGAGGCGGCCATGCGCACGTCCGCAAGCGCTTCAAGAGCCACGCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTCTTAAAGCATGTCAGAAATCAGA 2362  
2347 TACCTGTGACAGACCTCCAGCGGTACATGCGACAGTCTGTGCTCACTGAG-----GA 2400  
2363 TGCCAGTCACTGAGGAACCTCGTGTGTCATGAGCAGAGCATCTCTATGAATCAGAGCAG 2422  
2401 GACCAGCCGCTGAGGGATGCGTCTGTCATGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCTGTTGACTCTCTCTGCACTTCTGCTGTCACAGTGTGTAAGATTTGGTGA 2482  
2461 CAGTGGCTCTTCGACGCTCTCTCTACGCTTCTATGTCACACCGCGCTGCGCATCAGGG 2520  
2483 CAGGTGCTATACGAGTCCAGGGCATCCCCAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAACTCTACGTCAGTGCAGGGATCCCGAGGGTCTCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGGAGAACAGCTGTTGCTGAGGTGAGCGGATGGGTT 2602  
2581 CAGCCTGTGCTACGCGACATGGAGAACAGCTGTTGCGGGATTCGGCGGACCGGCT 2640  
2603 GCTTTTACGTTTGTGATGATTTCTGCTGACGCGCTCACTTGGACCAACCAAAAC 2662  
2641 GCTCTGCTGTTTGGTGGATGATTTCTGTTGGTACACTCACCTCACCGCGGAAAC 2700  
2663 CTTCTCAGCAGCCCTGGTCCATGGGTTCTCTGATGAGTGGGTGATATAAACTTCAGAA 2722  
2701 CTTCTCAGCAGCCCTGGTCCAGGTGTCCTGATGATGGTGGGTGATGCTTGGGAA 2760  
2723 GACAGTGTGAACTTCCCTGTGGAGCCCTGTTACCTGGTGGTGGTGCAGCTCCATCAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2820

QY 2783 GCCTGCTACTGCTCTTTCTCTGCTGGCTTGTGCTGGACACTCAGACTTTTGAGGT 2842  
Db 2821 GCGGGCCCAAGGCTTATCCCTGCTGGGCTGCTGCTGGATACCCGACCCCTGGAGGT 2880  
QY 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCAGCTTCCAGAG 2902  
Db 2881 GCAGAGGACTACTCCAGCTATGCCGAGCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
QY 2903 TGTCTTCAAAGCTGGGAAGACCATGCGGAACAAGCTCCTGCTGGTCTTGGGTTGAAGTG 2962  
Db 2941 CGGCTTCAAGCTGGGAGGAACATGCTCGCAACTCTTTTGGGGTCTTGGGCTGAAGTG 3000  
QY 2963 TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGATTTGAGGTGAACAGCTCCAGACGGTGTGCACCAACATCTA 3060  
QY 3023 CAAAGATCTCTCTGCTCAGCCTACAGGTTCAAGTTCATGATGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAAGATCTCTCTGCTCAGCCTACAGGTTTCAAGTTCATGATGTGATTCAGCTTCCCTTTCA 3120  
QY 3083 CCAGCGTGTAGGAAGAACTCACTCTCTTCTGGCATCATCTCCAGCAAGCATCCTG 3142  
Db 3121 TCAGCAAGTTTGAAGAACCCCACTTTTCTGGGCTCATCTCTGACGCGCTCCT 3180  
QY 3143 CTGCTATGCTATCTCTGAGGTCAAGATCCAGGAATGACACTAAAG-----GCCTC 3193  
Db 3181 CTGCTACTCATCTCTGAAAGCCAAAGACGAGGATGCTGCTGGGGCCCAAGGGCGCGC 3240  
QY 3194 TGCTCTCTTCTCTGAGCCGACATTTGGCTGCTGCTACGAGGCTTCTGCTCAAGCT 3253  
Db 3241 CGGCGCTCTCTCTGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
QY 3254 GGCTGCTCATCTCTGCTCATCTACAAATGTCTCTGGGACCTCTGAGCAGAGCCCAAACT 3313  
Db 3301 GACTCGACACCGTGTCACTACGTGCCACTCTCTGGGTCACTCAGGACAGCCAGACGCA 3360  
QY 3314 GCTGTGCGGAAGCTCCAGAGCGCAATGACCATCTTAAAGTGTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTGGGAAGCTCCGGGAGCAGCGTCTGCTGCTGGAGGGCGGAGCAACCCGCGC 3420  
QY 3374 CCTAAGCACAGACTTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCGCTCAGACTTCAGAGCCATCTGAGCTGA 3454

## RESULT 9

US-09-912-951-1  
; Sequence 1, Application US/08912951  
; Patent No. 6475789  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
; TITLE OF INVENTION: THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 335  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
OTHER INFORMATION: /note= "human telomerase reverse  
OTHER INFORMATION: transcriptase (hTERT) catalytic protein  
OTHER INFORMATION: component"

-08-912-951-1

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
23 CCGCGCTTTGAGCAAAATGACCCCGGCTTCTGCTGTTGCCCGCGTGGCTCTCTGCTGG 82  
40 CCGCGCAACCCCGGAGTGGCGGCTCCCGCTGGCGAGCGGTCCCTGCTGG 99  
83 CAGCGCATACCGGAGGTGTGCGCTGCGCACTTTGTGGCGGCTGGGGCCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGCTGCTGCGCACTTTGTGGCGGCTGGGGCCCGAGGG 159  
143 CAGCGCGCTTGTCAACCCCGGAGTGTGCGAGACTCTACCGCACTTTGTGTCCTAATGCT 202  
160 CTGGCGCTGTGTGAGCGCGGGGACCCCGCGCTTTCCGCGCTGTGGTGTGCGCTGTGCT 219  
203 AGTGTGATGCTGCGGCTCAGCGCTCCCGCTGCGGCTTTCCCTTCCACAGAGTGTG 262  
220 GGTGTGCTGCTTGGAGCGCAGCGCGCGCCCGCCCGCCCTCTCTTCCGCTGAGTGTG 279  
263 ATCCCTGAAGAGTGGTGGCCAGGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAAA 322  
280 CTGCTGAAGAGTGGTGGCCGAGGTGTGCGAGGCTGTGCGAGCGCGCGCGAGAA 339  
323 CGTGTGCTTTTGGCTTGTGCTCTTAACGAGCGCAGAGCGCGGCTTCCCATGGCTT 382

Db QY 340 CGTGTGCTGGCTTCGGCTTCGCGCTGCTGGAGCGGGGCCCGCGGGGGCCCCCGAGGCGCTT 399  
QY 383 CACTAGTAGCGTGGTAGCTACTTGGCCCAACTGTTATTAGAGACCTGCGTGTCACTGG 442  
Db 400 CACCACCGAGCGTGGCAGCTACTGCTCCCAACACGCTGACCGACGCACTGCGGGGAGCGG 459  
QY 443 TGCATGATGCTACTGTTGAGCGGAGTGGGCGACACCTGCTGGTCTACTGCTGTGGCACA 502  
Db 460 GGGTGGGGGCTGCTGCTGGCGCGGCTGGCGACGACGCTGCTGGTTCACTGCTGGCAG 519  
QY 503 CTGTGCTCTTTATCTTTCTGGTGGCCCCAGCTGTGCTACAGAGTGTGTGGTCTTCCCT 562  
Db 520 CTGCGGCTCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 579  
QY 563 GTACCAATTTGTGCCACACGATATCTGGGCTCTGTGTCCGTAGTTACAGGCCAC 622  
Db 580 GTACAGCTCGGCGCTGCCACTCAGGCCCGCGCCCGGACACAGCTAGTGGACCCGAG 639  
QY 623 CCGACCGCTGGGAGGAATTTCACTAACCTTAGTCTTTTACACAGATCAAGAGCAGTAG 682  
Db 640 GCGTCTGGGATGC-----GAACGGGCTTGGACCATAG 672  
QY 683 TCGCCAGGAGCACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
Db 673 CGTCAGGAGGCGCGGCTCCCTGGGCTTGCAGGCCCGGCTGCGAGGAGCGCGGGGG 732  
QY 743 TCTCACCAGTACAAGTGTGCTTTCAGCTAAGAGCGCAGATCTATCTCTGCTCCGAGAGT 802  
Db 733 CAGTGCCACCGGAGTCTGCGCTTGCCTCAAGAGGCGCGCTGCGGCTGCTCCCTGAGCC 792  
QY 803 GGAGGAGGAGCCCCCAGCAGGAGTGTCTACCAACCCCATCAGGCAATCATGGTGCCAG 862  
Db 793 GGAGCGGAGCCCCGTTGGGAGGCGGCTCTGGGCGCACCCCGGCGAGCGCGTGGACCGAG 852  
QY 863 TCCTGCTCGTCCCGGAGTGCCTACT-----GCAGAGAAAGATTTCCTCTCTAA 913  
Db 853 TGACCGTGGTTTCTGTGTGGTGTCACTGCCAGACCCCGGAGAGCCACCTCTTTGGA 912  
QY 914 AGGAAAGTGTCTGACCTGAGTCTCTC-----TGCGTGGTGTGCTGTAAACAAGCCAG 970  
Db 913 GGGTGGCTCTCTGGCAGCGCGCACTCCCAACCCATCCGTGGGCGCGCAGCACCGCGGG 972  
QY 971 CTCACAT---CTCTGCTGCACACCCCGCCAAATGCTTTTTCAGCTCAGGCACTTTAT 1027  
Db 973 CCCCCATCCACATCGCGGCGCACAGTCTCTGGGACAGCCTTGTCTCCCGGCTGTACGC 1032  
QY 1028 TGAGACAGACATTTCTTTTACTCCAGGGGAGATGCCAAGAGCGCTTAAACCCCTCAT 1087  
Db 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTCGGCGCTCTT 1089  
QY 1088 CTTACTCAGCAACCTCCAGCCTAACTTGTGCTGGGCGCAGGAGACTGGTGGAGATCATCTT 1147  
Db 1090 CTTACTCAGTCTCTGAGGCGCCAGCCTGACTTGGCGCTCGAGGCTCGTGGAGACCACTT 1149  
QY 1148 TCTGGGCTCAAGGCTAGGACATCAGGACCATCTCTCAGGACACACCGCTCTATCGGCTG 1207  
Db 1150 TCTGGTTCAGGCGCTTGGATGCCAGGACTCCCGCAGGCTTCCCGCGCTTCCCGCAGG 1209  
QY 1208 ATCTGGAGATGGGCGCCCTGTTCACAGCTGCTGTGTAACCATGACAGATGGCAATA 1267  
Db 1210 CTACTGGCAATGCGGCGCCCTGTTCTGGAGCTGTGTAACCATGACAGTGGCCCTTA 1269  
QY 1268 TGTGAGACTCTCTCAGGTCACTTTCAGGTTTCGAAACAGCAAAACCAACAGGTGACAGATGC 1327  
Db 1270 CGGGGTGCTCTCTCAGAGCGCACTGCTCCGCTGCGAGTGTCCCGTCAACCCAGGCGGTGT 1329  
QY 1328 CT-----TGAAACACGAGCC 1342  
Db 1330 CTGTGCGCGGAGAGCCCGAGGCTCTGTGGCGGCGCCCGAGGAGGAGGACACAGACC 1389  
QY 1343 ACCGCACTCATGATTTGCTCGCGCTGACAGCAGTCTCCCTGGCAGGTATATGTTTTCT 1402  
Db 1390 CCGTGGCTGGTGCAGCTGTCTCGCCAGACACAGAGCCCTTGGCAGGTGTACGGCTCT 1449

1403 TCGGGCTGTCTCTGGAAGTGGTGTCTGTAGTCTCTGGGGTACGAGCACAATGAGCG 1462  
1450 GCGGGCTGTCTCTGCGCGCTGTGGTGGCCCCCAGGGCTCTGGGGTCCAGGACAAGAAAG 1509  
1463 CCGCTCTTTTAAGAACTTAAGAGTTTCACTCTGTTGGGGAATACGGAAGTATCACT 1522  
1510 CCGCTCTCTAGGAACCAAGAGTTTCACTCTCCCTGGGGAAGCATGCCAAGCTCTCGCT 1569  
1523 GCAGGAAGTGTGGAGATGAAGTAGAGGATGACCTGCTCGGAGAGCCGCGG 1582  
1570 GCAGGAGCTGACGTGGAGATGACGCTGCGGGAAGTCTGCTGCTCGGAGAGCCGAG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGGACCGCTCTGAGGAGAGGATCCTGGCTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGCGCGCAGAGACCGCTCTGCGTGGAGATCCTGGCCAAAGTTCT 1689  
1643 GTTCTGGTGTAGACACATACGTGGTACAGCTGTAGTGTATCTTTTATCATCACAAG 1702  
1690 GCACTGGCTGATGATGTGTACGTCGTGCGAGCTGCTCAGGTCTTTCTTTATGTACGGA 1749  
1703 GAGCACATCCAGAGAACAGGCTCTTTCTTACCGTAAAGTGTGTGAGCAAGTGA 1762  
1750 GACCAAGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGA 1809  
1763 GAGCATTTGGAGTACAGCAACCTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGAGA 1822  
1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTGCGAAGCAGA 1869  
1823 GGTGAGGATCACCAGACAGCTGGCTAGGACATGCCATCTGCAGACTCGGTTTATCCC 1882  
1870 GGTGAGGAGATCGGGAAGCCAGGCGCGGCTGCTGAGCTCCAGACTCCGTTTATCCC 1929  
1883 CAAGCCCAAGCGCTGCGGCGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTTGAAGGCTGCGGCGGATTTGAACATGAGTACGTCTGGTGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGCAGCCAGCATTTTCAACCAAGCTGTCAAGACTCTCTTCAAGATGCT 2002  
1990 CCGCAGAGAAAGAGGCGGAGGCTCTCACTCGAGGGTGAAGGCACTGTTCAAGGCTGT 2049  
2003 CAATATGAGCGGACAAACATCTCACTTATGAGGCTCTTGTACTGGGTATGAATGA 2062  
2050 CAATCAAGAGCGGCGGCGCGCGGCTCTTGGGCGCTCTGTGCTGGGCTGGACGA 2109  
2063 CATCTACAGACCTGGCGGCTTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2122  
2110 TATCCACAGGCGCTGGCGGACCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2169  
2123 GATGTATTTTGAAGCAGATGTGACGGGCGCTATGATGCCATCCCGGCTAAGCT 2182  
2170 GCTGTACTTTTGAAGTGTGTGACGGGCGGTACGACACCATCCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTGTTGCCAATATGATCAGGCACTCGGAGACAGTACTGTATCCGCGAGTA 2242  
2230 CACGAGGCTCATCGCCAGCATCA---AAGCCAGAACAGTACTGTGCTGGTGGT 2286  
2243 TGCAGTGTCCGAGATAGCCAGCCAGTTCACAGTCTCTTTAGAGACAGGTAC 2302  
2287 TGCCTGTTCAGAGGCGCGGATGAGGCGGCTGCGGACGCTCGCAAGGCTTCAAGAGCCAGTCTC 2346  
2303 CACCTCTCTGACCTCCAGCATACATGAGGCGGCTTCTTAAAGCATCTCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGTGCTAGGAACTCGGTTGTGTCATGAGCAGAGCATCTCTATGAATGAGAGAG 2422  
2401 GACCAGCCCGCTGAGGGATGCCCTCGTATCATGAGCAGAGCTCTCTCCCTGAATGAGCCAG 2460  
2423 CAGCAGCTGTTTCACTCTCTCTGCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCAGCTCTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520

2483 CAGGTGCTATACGAGTCCAGGGCATCCCGAGGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAGTCTCTAGTCCAGTCCAGGGATCCCGAGGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAAGCTGTTTGTGAGGTGACGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGGGACATGGAGAAAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTACGTTTGTGTGATGACTTCTGTGTGTGAGCTCTACCTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCACCCACGCGAAAC 2700  
2663 CTTCTCTCAGCACCTGCTCCATGGGCTTCTGAGTATGGGTGCATATAAACTTCAGAA 2722  
2701 CTTCTCTCAGGACCTGCTCGAGGTGTCCCTGAGTATGGCTGGTGGTGAATTCGGAA 2760  
2723 GACAGTGTGAATTCCTCTGTGAGGCTGTATACCTGGGTGTGAGTCCATACCACT 2782  
2761 GACAGTGTGAATTCCTCTGTAGAAGACGAGGCCCTGGGTGCGACGGCTTTTGTTCAGT 2820  
2783 GCTGTCTCACTGCTGCTTCTCTGCTGTGGTGTGCTGTGGACACTCAGACTTTTGGAGT 2842  
2821 GCGGCGCCAGCGGCTATTCCTCTGTGTGGGCTGTGCTGTGATACCCGAGCCCTGGAGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACCTTCCAGAG 2902  
2881 GCAGAGCGACTACTCCAGCTATGCCGAGACCTCCATCAGAGCAGTCTCACTTCAACCG 2940  
2903 TGTCTTCAAAGCTGGGAGACCATGGGAAAGAGTCTCTGCTGTGGTTCGGTTCGGTGGAGT 2962  
2941 CGGCTTCAAAGCTGGGAGGAAAGTGGTGGGAACTCTTGGGGTCTTGGGGCTGGAAGT 3000  
2963 TCAGGCTGTATTTTACAGTTCAGGTGAAGAGCCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCAGAGCTGTTTCTGATTTGAGTGAAGAGCCTCCAGAGCTGTGCACCAACATCTA 3060  
3023 CAAGATCTCTGCTTCCAGGCTACAGTTCAATGATGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAGATCTCTGCTGCGAGGCTACAGTTTCAAGCTGTGCTGAGCTCCCATTTCA 3120  
3083 CCAGCGTGTAGGAAGACCTCACATTTCTTCTGGGCTCATCTCCAGCCAGCATCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTGACACGGCTCCCT 3180  
3143 CTGCTATGCTATCTGAAGGTCAAGATCCAGGAATGACATAAG-----GCTC 3193  
3181 CTGCTACTCCATCTGAAAGCCAGAACGAGGATGTGCTGGGGCCAGAGGGCGCGC 3240  
3194 TGGCTCTTCTCTCTGAGCGGACATTTGGCTGTCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTGCTCTGAGGCGGTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 3300  
3254 GGTGCTCTTCTGCTATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTCGACACCGTGTCACTAGTCCACTCTGCGGTCTCTCAGGACAGCCCAAGAGCA 3360  
3314 GCTGTCCGGAAGCTCCAGAGCCAGAAAGACCATCTTAAAGCTGAGTGAAGCTGAGC 3373  
3361 GCTGAGTCCGAAGCTCCCGGAGCAGCTGACTGCTGCTGAGGCGGAGCAACCCGCGC 3420  
3374 CTTAAGCAGACACTTTCAGACCATTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTCAGACCATCTCTGACTGA 3454

RESULT 10  
US-09-733-294A-3  
; Sequence 3, Application US/09733294A  
; Patent No. 6492171  
; GENERAL INFORMATION:  
; APPLICANT: Brett P. Monia  
; APPLICANT: William Gaarde  
; APPLICANT: Susan M. Freier

APPLICANT: Edward V. Wanciewicz  
TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
FILE REFERENCE: ISPH-0527

CURRENT APPLICATION NUMBER: US/09/733,294A

CURRENT FILING DATE: 2000-12-07

PRIOR APPLICATION NUMBER: 09/572,423

PRIOR FILING DATE: 2000-05-16

NUMBER OF SEQ ID NOS: 108

SEQ ID NO 3

LENGTH: 4015

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (56)...(3454)

09-733-294A-3

Query Match 45.8%; Score 1600.8; DB 4; Length 4015;

Best Local Similarity 68.9%; Pred. No. 0;

Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

23 CCCGCGCTTGAGCAATGACCCCGGCTCTCGTTGCGCCCGCGGTGCGTCTCTGTCGG 82

40 CCCGCGCGCCCGCGATGCGCGCGCTCCCGCTGCCGAGCGGTGCGTCTCTGTCGG 99

83 CAGCGGATACCGGAGGTGTGCGCGCTGGCAACCTTTGTGCGCGCGCTGCGGCGCCGAGGG 142

100 CAGCCACTACCGGAGGTGTGCGCGCTGGCCACGTTGTTGCGCGCGCTGCGGCGCCGAGGG 159

143 CAGCGCGCTTGTGMAACCCGGGAGCCGAGATCTACCGACTTTGGTGGCCCAATGCTT 202

160 CTGGCGGCTGTGACGCGCGGGACCCGGCGGCTTTCCGCGCGCTGGTGGCCCGAGTGCCT 219

203 AGTGTGCATGCACTGGGCGCTCAGACGCTCACCCTGCGGACCTTTCTTCTCCACAGGTGC 262

220 GGTGTGCTTGGGACGACGCGCGCGCCCGCGCGCTCTCTTCCGTCAGGTGC 279

263 ATCCCTGAAGAGCTGTGGCGAGGTGTGACAGACTCTGTGAGCGCAACGAGAGAA 322

280 CTGGCTGAAGAGCTGTGGCGCGAGGTGTGACAGGCTGTGAGCGCGCGCGCGCGAGAA 339

323 CTGTGTGCTTTTGGCTTTGAGCTGTGTAAGCGGCCAGAGGGCGGCGCTCCCATGCCCTT 382

340 CGTGTGCGCTTTCGCTTCGCGCTGTGAGCGGGCCCGCGGGGCGCCCGCGAGGCTT 399

383 CACTAGTAGCGTGTGCTACTTTCGCCAACACTGTATTAGAACCTCGGTGTGAGTGG 442

400 CACCAACAGCGTGCAGCTACTTCCCGCAACACACGCTGACCGACGCACTGGCGGGAGCGG 459

443 TGCATGGATGCTACTGTTGAGCGAGTGGCGGACGACCTGCTGTCTACCTGTGCGCACA 502

460 GCGGTGGGCGTGTGCTGCGCGCGGTGGCGGACGACGTGCTGTTCACTGTCTGCGACG 519

503 CTGTGCTCTTTATCTTCTGTGTCGCCCGCGAGTGTGCTTACCGAGTGTGGGTCTCCCGCT 562

520 CTGCGCGCTCTTTGTGCTGTGGTCTCCCGAGTGTGCGCTTACCGAGTGTGGCGCGCGCT 579

563 GTACCAATTTGTGCCACACGAGATCTGCGCTCTGTGTCGCTAGTTACAGGCGCCAC 622

580 GTACGAGTTCGCGCTGTGCACTCAGGCGCGCGCGCGCGCACACGCTAGTGGAC 639

623 CCGACCGGTGGCGAGGAATTTCACTAAGCTTAGGTCTTTACACAGATCAAGAGCAGTAG 682

640 GCGTCTGGGATGC-----GACCGGCGCTGGAAACCATAG 672

683 TGCACGAGAGCACCGAACCCCTGGCGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742

673 CGTCAAGGAGGCGCGGCTCCCGCTGGCGCTCCACGCGCGGTCGAGAGGCGCGCGGG 732

743 TCTCACAGTACAAAGTGTGCTTACGTTAAGAGCCAGATGCTATCTCTGCCGAGAGT 802

733 CAGTGCACCGGAGTCTGCGCTTGTGCCCAAGAGGCGCGGCGTGGCGCTGCCCTTGAGCC 792

QY 803 GGAGAGGGACCCCAAGCAGGTGCTTACCAACCCCATCAGCAATCATGGTGCCCAAG 862

Db 793 GGAGCGGACCGCGTTGGCGCAGGGTCTTGGGCCCAACCCGGGAGGAGCGGTGACCGAG 852

QY 863 TCCTGTCTCGGTCCCGCGAGGTGCTTACT-----GCAGAGAAAGATTTGTCTTTAA 913

Db 853 TGACCGTGGTTTCTGTGTGTGTCTACCTGCCAGACCCCGGAGGAGCCACTCTCTTTGGA 912

QY 914 AGAAAGGTGTGACCTGAGTCTCTC---TGGGTGGTGTGCTGTAAACACAGCCCGAG 970

Db 913 GGGTGGCTCTCTGGCACGCGCACTCCACCACTCGTGGGCGCGGAGCCAGCCGCGGG 972

QY 971 CTCACAT---CTCTGTCTCACACCCCGCCAAATGCTTTTCAGCTCAGGCCCATTTAT 1027

Db 973 CCCCCATCCACATCGGGCCACCAAGTCCCTGGGACACGCTTGTCTCCCGGTATCGC 1032

QY 1028 TGAGACGAGACATTCCTTTACTTCCAGGGAGATGGCCAGAGCGTCTTAAACCCCTCAT 1087

Db 1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCGCTCTT 1089

QY 1088 CTTACTCAGCAACTCTCAGCTTAACTTGTGCGGCGAGGAGACTGGTGGAGATCATCTT 1147

Db 1090 CTTACTCAGCTCTCTCAGGCGCGAGCTGACTGGCGCTCGAGGCTCTGTGGAGCATCTT 1149

QY 1148 TGTGGGTCAAGGCTTAGACATCAGGACCACTTGTGAGGACACACCGTCTATCGCGTCG 1207

Db 1150 TCTGGGTTCCAGGCGCTGGATGCCAGGACTCCCGCAGGTTGCGCCGCTTGCCTCCAGCG 1209

QY 1208 ATACTGCGACATCGGCGCTCTTCCAACTGTGTTGAACTGAGAGTGCAGGAGTGCATA 1267

Db 1210 CTTACTGCGCAATGCGCGCTCTTTCTGAGGTGCTTGGAGACCACTGCGGAGTGCCTTA 1269

QY 1268 TGTGAGCTCCTCAGGTCACTTGTGAGGTTTGGAAAGCAACCAAGAGGTGACAGATGC 1327

Db 1270 CGGGGTCTCTCAAGACGCACTGCGCGCTGCGAGTGTGCGGTCACTCCCGCAGCGCGTGT 1329

QY 1328 CT-----TGAAACACCGAGCC 1342

Db 1330 CTGTGCGCGGAGAACCGCCAGGCTCTGTGGGCGCCCGGAGGAGGAGACACAGACCC 1389

QY 1343 ACCGACCTCATGATTTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTTCT 1402

Db 1390 CCGTGCCTCGTGTGCTGCTCGCGCAGCACAGCAGCCCTGGCAGGTGTACCGCTTCGT 1449

QY 1403 TCGGCGCTCTCTGCAAGGTGTGCTGCTAGTCTTGGGTGTACAGGCAATGAGCG 1462

Db 1450 GCGGCGCTGCTGCGCGCGTGTGCGCCCGCGCTTGGGCTCCAGGCAACAGAACG 1509

QY 1463 CCGCTTCTTTAAGAACTTAAAGAGTTTCACTCTGTGGGAAATACGGCAAGCTATCACT 1522

Db 1510 CCGCTTCTCAGGAACACCAAGAGTTTCACTCTCCCTGGGGAAGCATGCAAGCTCTCGCT 1569

QY 1523 GCAGGAATGATGTGGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGCAGCGCGGG 1582

Db 1570 GCAGGAGCTGACGTGAAGATGAGCGTGCAGGACTGCGCTGGCGGAGGAGGCGCGG 1629

QY 1583 GAAGGACCGTGTCCCGCTGCAGACACCGTCTGAGGGAGAGGATCCTGGCTACCTTCTCT 1642

Db 1630 GGTGGCTGTGTTCCGCGCGCAGAGCACGCTGTGCTGAGGAGATCTCTGGCCAGTTCT 1689

QY 1643 GTTCTGCTGATGAGACATACGTTGATCAGCTGTGTTAGGTCACTTTTATCATCAGAGA 1702

Db 1690 GCATGCTGATGAGTGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCTACGGA 1749

QY 1703 GAGCACATTCAGGAAGAACAGGCTCTTCTTACCGTAAGAGTGTGTGAGCAAGTGTCA 1762

Db 1750 GACCACTTTTCAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGCAAGTTGCA 1809

QY 1763 GAGCATTTGAGTCCAGCAACCTTGAAGAGTGTGCGGTACGCGAGCTCTCAAGAGGA 1822

Db 1810 AAGCATTTGGAATCAGACAGCACTTGAAGAGGTGTGAGTGTGCGGAGCTGTGGAAGCAGA 1869

QY 1823 GGTGAGGATCAACAGACACCTGGTGTAGCCATGCGCCATCTGCAGACTGCGCTTCACTCC 1882

1870 GGTGAGGAGGATCGGGAAGCCAGCGCCCGCTGTCAGCTCCAGACTCCGCTTATCC 1929  
1883 CAAGCCCAACCGCCCTGCGGCCCACTTGTGAACATGAGTTATAGCATGGTACAGAGCTTT 1942  
1930 CAAGCTGACGGGTGCGGCCGATTTGTGAACATGAGTTATAGCATGGTACAGAGCTTT 1989  
1943 GGGCAGAGGAGCAGCGCCCGCCAGCTTTCAGCCAGCGTCTCAAGACTCTTTCAGCATGCT 2002  
1950 CCGCAGAGAAAGAGCGCCCGAGCTCTCACTCGAGGGTGAAGGCACCTGTTGAGGGTCT 2049  
2003 CAACCTATGAGCGGCAAAAACATCTCACTTATGGGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAACCTACGAGCGGGCGCGGCCCGCCCTCTGGGGCTCTGTGCTGGGCTGAGCGA 2109  
2063 CATCTAAGAGCCTGGCGGCCCTTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2122  
2110 TATCCAGAGGCGCTGGGCGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169  
2123 GATGTACTTGTGTAAGGCAGATGTGACCGGGGCTATGATGCCATCCCGCAGGGTAAGCT 2182  
2170 GCTGTACTTGTCAAGGTGATGTGACGGGGGCTAGCAGACCATCCCGCAGGAGGCT 2229  
2183 GGTGAGGTTGTTCCAAATATGATCAGGCACTCGGAGAGCAGTACTGTATCGCCAGTA 2242  
2230 CACGAGGTCATCGCCAGCATCA---AACCCAGAACACCTACTGCGTGGTGGTGA 2286  
2243 TGCAGTGTCCGAGAGATACCCAGGCGCAGTCCCAAGTCCCTTTAGGAGCAGGTAC 2302  
2287 TGCCGTGTCCAGAGGCGGCCCATGCGGCGCTCGGCAAGGCTTCAAGAGCCACGTCTC 2346  
2303 CACCCCTCTGACCTCCAGCCATACATGCGCCAGTCTCTTAAGCATCTGCGAGGATTCAG 2362  
2347 TACCTTGACAGACCTCCAGCGGTATACGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCATGAGAACTCCGTTGTCATCGAGCAGAGCATCTATGAATGAGAGCG 2422  
2401 GACAGCCCGCTGAGGAGTGCCTGCTCATGAGCAGAGCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCCTGTTGACTTCTTCTGCACTTCTCTGCTGTCAGAGTGTGTAAGATTGGTGA 2482  
2461 CAGTGGCTCTTCCAGCTCTTCTTCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGTGTCTATGAGCAGTGCAGGAGCTCCCGCCAGGCTCCAGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACCTGAGTGCAGGAGCTCCCGCAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAGAACAGCTGTTGCTGAGTGCAGCGGATGGGTT 2602  
2581 CAGCTGTGCTACGGGACATGAGAGAACAGCTGTTTGGGGATTCGGCGGACGGGT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTTGTGAGCCCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTCGGTTTGTGATGATTTCTTGTGAGCTTCTTGTGAGCTTCTGAGTTCGGGAA 2700  
2663 CTTCCTCAGCAGCTGCTCCATGCGCTTCTGATGATGCGTGCATGAATTAAGTCCAGAA 2722  
2701 CTTCCTCAGGACCTGCTGAGGTGCTGAGTGTCCCTGAGTATGGTGGTGAATTCGGAA 2760  
2723 GACAGTGTGAACCTTCCCTGTTGAGCCTGTTACCTGGGTGCTGAGCTCCATACAGCT 2782  
2761 GACAGTGTGAACCTTCCCTGTTGAGAGCAGGCGCTGGTGGCAGCGCTTGTTCAGAT 2820  
2783 GCTGCTCAGTCTGTTTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2842  
2821 GCGGCGCCAGCGGCTATTCCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACTCAATTAAGACGAGCTCACTTCCAGAG 2902  
2881 GCAGAGCAGCTACTTCCAGTATGCCGAGCTCCATCAGAGCAGCTCACTTCAACCG 2940  
2903 TGTCTTAAAGCTGGAGAGCCATGCGGAAACAAGCTCCTGCTGCTGCTGCTGCTGCTGCTG 2962

Db 2941 CGGCTTCAAGGCTGGAGGAACATGGCTCGAAACTCTTTGGGGTCTTGGGCTGAAGT 3000  
Qy 2963 TCACGGTCTATTTCTAGACTTCAGGTGAACAGCTTCAGAGAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGGCTGTTTCTGATTTGAGGTGAACAGCTTCAGAGCGTGTGCACCAACATCTA 3060  
Qy 3023 CAAGATCTTCTGCTTCAGGCTTCAGGTTCCATCATGTGTGATTCAGCTTCCCTTTGA 3082  
Db 3061 CAAGATCTTCTGCTTCAGGCTTCAGGTTTCAGCATGTGTGCTGAGCTCCCAATTTCA 3120  
Qy 3083 CCAGCGTGTGTAGGAAGACTCACAATTTCTTCTGGGCTCATCTCCAGCCAGCATCTG 3142  
Db 3121 TCAGCAAGTTTGAAGAAACCCCAATTTTCTGCGGTCACTCTGACACGGCCTCCCT 3180  
Qy 3143 CTGCTATGCTATCTTCAAGGTCAAGATCCAGGAATGACACTAAG-----GCCTC 3193  
Db 3181 CTGCTACTCCATCTTAAAGCCAGAACCGAGGATGTGCTGGGGGCAAGGGCGCCGC 3240  
Qy 3194 TGGCTCTCTTCTTCTTCAAGCGGCAATTTGGCTCTGTACCAAGGCTTCTTCTCTCAAGCT 3253  
Db 3241 CGGCTCTCTGCTTCCAGGCGGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
Qy 3254 GGTGCTCATTTCTGCTCATCTACAATGTCTCTGGGACCTCTGAGGACAGCCCAAACT 3313  
Db 3301 GACTCGACACGCTGTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
Qy 3314 GCTGTGCGGAGCTCCAGAGCGCAATGACCATCTTAAAGCTGAGCTGAGCCAGC 3373  
Db 3361 GCTGAGTGGAGCTTCCGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
Qy 3374 CTTAAGCAGACACTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCTCTGAGCTTCAAGACCATCTGAGCTGA 3454

## RESULT 11

US-08-974-549A-343  
; Sequence 343, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 343:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4037 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /note= "refined sequence of hTERT cDNA"  
-08-974-549A-343

Query Match 45.7%; Score 1596; DB 3; Length 4037;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 2376; Conservative 0; Mismatches 970; Indels 108; Gaps 9;  
23 CCGGGCTTTGACCAATGACCCGGCTCTCGTTGCCCCCGGTGGCGCTCCCTGCTGG 82  
40 CCGGGCCACCCCGCGATGCGCGGCTCCCGCTGCGGACCGCTCCCTGCTGG 99  
83 CAGCCGATACCGGAGGTGTGGCCGCTGGCAACCTTTGTGGCGGCTGGGGCCGAGGG 142  
100 CAGCCACTACCGGAGGTGTGGCCGCTGGCAACCTTTGTGGCGGCTGGGGCCGAGGG 159  
143 CAGCGCGCTTTGCAACCCGGGACCCGGAAGATCTACCGCACTTTGGTTGCCCAATGCCCT 202  
160 CTGGCGCTGTGTGACGCGCGGGAACCGCGGCTTTCCGCGGCTGTGTGGCCAGTGCCT 219  
203 AGTGTGATGCACTGGGGTTCACGCTTCACCTGCGCACTTTCCCTTTCACAGAGTGC 262  
220 GGTGTGCGTGGCTGGGAGCGACCGCGCCCGCCCGCCCTCCCTTCCCGCAGGTGC 279  
263 ATCCCTGAAGAGCTGTGGCCAGGTTGTGAGAGCTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCTGAAGAGCTGTGGCCCGAGTGTGTGAGAGCTGTGTGAGCGCGGCGGAGAA 339  
323 CGTGTGCTTTTGGCTTGTAGCTCTTAAAGAGCCAGAGCGGGGCTCCCAATGGCTT 382  
340 CGTGTGCGCTTGGCTTGTGAGCTGTGAGAGCGCGGCGGCGGCGGCGGCGGCTT 399  
383 CACTAGTACGCTGCTAGTACTTCCCAACACTGTATTGAGACCTGCGGTGTCAGTGG 442  
400 CACCACCGGTGCGGAGCTACTCTCCCAACACGCTGACCGCACTGCGGGGAGCGG 459

QY 443 TGCATGATGCTACTGTTGAGCGAGTGGCGACGACCTGTGTTCTACTCTGCTGGCACA 502  
DB 460 GGGTGGGGCTGCTGCTGGCGCGCTGGCGACGACGTCGTTGTTCACTCTGCTGGCAG 519  
QY 503 CTGTGCTCTTTATCTTCTGTGTCGCCCGCCAGCTGTGCTACCAAGGTGTGTGGGTCTCCCT 562  
DB 520 CTGGCGCTCTTTGTGCTGGTCTCCAGCTTGGCTTACAGGTGTGGGGCCGCGCT 579  
QY 563 GTACCAAAATTTGTGCCACCAAGATATCTGGCCCTCTGTGTCGCTAGTTACAGGCCAC 622  
DB 580 GTACCAAGCTCGGCGCTGCCACTCAGCCCGGCGCCCGCCAGCTAGTGGACCCGAG 639  
QY 623 CCGACCGTGGCGAGGAATTTCACTAACTTAGTCTTACAAAGATCAAGAGCAGTAG 682  
DB 640 GCGTCTGGGATGC-----GACGGGCTGGAACCATAG 672  
QY 683 TCGCCAGGAAGCACCGAAACCCCTGCTGCTTGCATCTCGAGGTACAAGAGCATCTGAG 742  
DB 673 CGTCAGGGAGCGGGGCTCCCTGCGCTTGCAGCCCGGCTGGAGAGCGGGGG 732  
QY 743 TCTCACCAGTACAGTGTGCTTCACTAAGAGGCGAGATCTATCTCTCCGAGAGT 802  
DB 733 CAGTGCCAGCGGAAGTCTGCGCTTGCAGAGGCGCGCTGGCGCTGCCCCCTGAGCC 792  
QY 803 GGAGGAGGAGCCCGCACAGGCGAGTGTCTACCAACCCCATCAGGCAATCATAGGTGCCAAG 862  
DB 793 GGAGCGGACGCCCTTGGGCGAGGGTCTTGGGCGCACCGCGGCGAGCGCTGGACCGAG 852  
QY 863 TCCTGTCTGCTCCCGGAGGTGCTACT-----GCAGAGAAAGATTTCTCTTAA 913  
DB 853 TGACCGTGGTTTCTGTGTGTGTCACCTGCGAGACCGCGGAGGAGGACACCTCTTTGA 912  
QY 914 AGGAAAGGTGCTGACCTGAGTCTCTC---TGGGTGCGTGTGCTGTAAACACAAGCCAG 970  
DB 913 GGGTGGCTCTCTGGCAGCGCCACTCCACCATCTCGTGGGCGCGCAGCACCGCGGG 972  
QY 971 CTCACAT---CTCTGTGTGTCACACCGCGCAAAATGCTTTTCTGAGTCTAGGCCATTTAT 1027  
DB 973 CCCCCCATCCACATCGCGGCCACACGCTCCCTGGGACAGCGCTTGTCCCCCGGTGTACGC 1032  
QY 1028 TGAGCAGACATTTCTTTTACTTCCAGGGAGATGGCCAGAGCGTCTAAACCCCTCAT 1087  
DB 1033 CGAGACCAAGCACTTCTCTTCTCTCTCAGCGA---CAAGGAGCAGCTCGGGCCCTCTT 1089  
QY 1088 CTTACTCAGAACTCCAGCTAACTGACTGGGGCAGAGACTGGTGGAGATCATCTT 1147  
DB 1090 CTTACTCAGCTCTCTGAGGCGCAGCCTGACTGGCGCTCGGAGGCTCGTGGAGACCATCTT 1149  
QY 1148 TCTGGGCTCAGGCTAGGACATCAGGACCATCTCTCAGGACACACCGTCTATCGGCTCG 1207  
DB 1150 TCTGGTTCCAGGCTTGGATCGAGGACTCCCGCAGTTGCCCGCTGCCCCAGCG 1209  
QY 1208 ATACTGCGATGCGGCCCTTGTTCACAGCTGTGTGTAACCATTCAGAGGTGCCAATA 1267  
DB 1210 CTACTGCAATGCGGCCCTGTTTCTGAGCTGTCTGGGAACCAACCGCAGTGGCCCTA 1269  
QY 1268 TGTGAGACTCTCAGTTCATTTGAGGTTTGGAAAGCAAAACCAAGCTGACAGATGC 1327  
DB 1270 CGGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGTGTGCGGTACCCCGAGCGCGGT 1329  
QY 1328 CT-----TGAAACACAGCC 1342  
DB 1330 CTGTGCCCGGGAAGACCCAGGGCTCTGTGGCGGCCCGCAGGAGGAGGACACAGACCC 1389  
QY 1343 ACCGACCTCATGATTTGCTCGGCTGTCAGCAGTCCCTGCGAGGTATATGTTTCT 1402  
DB 1390 CCGTGGCTGTGTGAGTGTCTCGCGCAGCACAGCAGCCCTGCGAGGTGTACGGCTTGT 1449  
QY 1403 TCGGCTGTCTCTGCAAGGTGTGTCTGCTAGTCTCTGGGGTACCCAGGCAAAATGAGCG 1462  
DB 1450 CCGGGCTGTCTGCGCGGCTGTGTGCCCCCAGGCTCTGGGGCTCCAGGCAACAAGAG 1509



1463 CCGCTTCTTTAAGAACTTAAGAGTTTCATCTCGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTTAGGAAACACCAAGAAAGTTCACTCTCTGGGAGACATGCCAAGCTCTCGCT 1569  
1523 GCAGAACTGATGTGGAAAGTAGAAGTAGAGGATGCCACTGCTCCGAGCAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAAGTAGAGCTGCGGAGTGCCTTGGCTCGCAGGAGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTCGAGAGCACCTCTGAGGAGAGAGATCCTGGCTACGTTCT 1642  
1630 GGTGGCTGTGTTCCGCGCGAGAGACCGCTCTGCGTAGAGAGATCCTGCCAAGTTCT 1689  
1643 GTTCTGCTGATGACACATACGTGGTACAGCTGCTTAGGTCACTCTTTTACATCAGAGA 1702  
1690 GCACTGGCTGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1749  
1703 GAGCACATTCAGAGAACAGGCTCTCTCTACCGTAAAGAGTGTGAGAGCAAGCTGCA 1762  
1750 GACCACGTTTCAAAGAACAGGCTCTTTTCTACCGCGCGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCAATTGGAGTCAGGCAACACCTTGAAGAGTGTGCGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCAATTGGAATCAGACAGCACTTGAAGAGGTGACAGTCCGCGAGCTGTGGAAGCAGA 1869  
1823 GGTGAGCATCACAGACACCTGCTGCTAGCCATGCGCATCTGAGACTGCGCTTATCCC 1882  
1870 GGTGAGCAGCATCGGGAAGCCAGGCGCGCTGCTGACGTCCAGACTCCGCTTATCCC 1929  
1883 CAAGCCCAAGCGCTGCGGCGCCATTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCCTGACGGGCTGCGGCGGATTTGAACATGAGTACGTCGTGGAGCAGAACGTT 1989  
1943 GGGCAGAGAGAGCAGGCGCCAGCATTTACCCAGCGTCTCAAGACTCTCTTCAAGATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGGCTCTCACTCGAGGGTGAAGGCACTGTTCAAGCGTCT 2049  
2003 CAATATGAGCGGCAAAACATCTCACTTATGAGGCTTCTGCTGCTGGGTATCAATGA 2062  
2050 CAATAGAGGCGGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2109  
2063 CATCTAGAGCCTGCGGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122  
2110 TATCCAGAGGCGCTGCGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169  
2123 GATGATCTTTTAAAGCAGATGACCGGGGCTATGATGATGATGATGATGATGATGATG 2182  
2170 GCTGATCTTTTCAAGGTGATGTGAGCGGCGGTACGACACATCCCGCCAGGAGGCT 2229  
2183 GGTGGAGGTTTGTGCCAATATGATCAGGCACCTCGAGAGCAGGTACTGTATCCGCCAGTA 2242  
2230 CACGGAGGTCACTGCCAGCATCATCA--AACCCAGAACACGTACTGCTGCTGCTGCTG 2286  
2243 TGCAAGTGTGCGGAGATAGCCAGGCAAGTCCACAGTCTCTTGAAGAGCAGGTCAAC 2302  
2287 TGCGGTGTCCAGAAAGCGCGCCATGCGGACGTCGCCAAGGCTCTTAAAGAGCAGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATAGGCGGAGTTCCTTAAAGCATCTCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGCTGCTGCTGCTGCTGCTG 2400  
2363 TGCAGTGTGCTAGGAACTCCGTTGTATCGAGAGCAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGAGTCCGTCGTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTTTGTGACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCGAGCTCTCTCTAAGCTTATGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGTGTCTATACGAGTGTCCAGGCAATGCCAGGCTCCAGCTTATCCACCTCTCTG 2542  
2521 CAAGTCTCTAGTCCAGTGTCCAGGAGTCCCGAGGCTCCATCTCTCTCCAGCTGCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAGAACAGCTGTTGCTGAGTGTCCAGCGGATGGGTT 2602

Db 2581 CAGCCTGTGCTACGGGCAATGGAGAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
Qy 2603 GCTTTTACGTTTGTGATGACTTTCTGTTGTGAGCGCTCACTTTGGACCAAGCAAAAC 2652  
Db 2641 GCTCTGCGTTTGGTGAATGTTCTTGTGTTGAGACCTCACCTCACCCACGCGAAAC 2700  
Qy 2663 CTTCTCTCAGACCTCGTCCATCGGTTTCTGAGTATGGTGCATGATAAACTTGCAGAA 2722  
Db 2701 CTTCTCTCAGACCTCGTCCGAGGTGCTCCTGAGTATGCTGCTGCTGCTGCTGCTG 2760  
Qy 2723 GACAGTGTGAATCTTCCCTGTGAGAGCTGTACCTCGGTGTGAGTGTGAGTGTGAGTGT 2782  
Db 2761 GACAGTGTGAATCTTCCCTGTGAGAGAGAGAGGCGCTGCTGCTGCTGCTGCTGCTG 2820  
Qy 2783 GCTGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
Db 2821 GCGGCGCCACGCGCTTATCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
Qy 2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCAGCTTCCAGAG 2902  
Db 2881 GCAGAGGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTGTCACTTCAACCG 2940  
Qy 2903 TGTCTTCAAAGCTGGAGAGACCATGCGGAAACAGCTCCTGCTGCTGCTGCTGCTGCTGCTG 2962  
Db 2941 CGCTTTCAAGGCTGGAGGAAACATGCTGCGAACTCTTTGGGGTCTTTGGGCTGAAAGT 3000  
Qy 2963 TCACGGCTGTTTCTGAGCTTGCAGGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3022  
Db 3001 TCACAGCTGTTTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 3060  
Qy 3023 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3082  
Db 3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
Qy 3083 CCAGCTGTTTGAAGAACCTCACAATCTTCTGCGCATCATCTCCAGCCAGCATCTCG 3142  
Db 3121 TCAGCAAGTTTGAAGAACCCCAATTTTCTGCGCTCATCTCTGACCGGCTCTCCT 3180  
Qy 3143 CTGCTATGCTATCTTGAAGGTCAAGATCCAGGAATGACATAAG-----GCTC 3193  
Db 3181 CTGCTACTCCATCTGAAAGCCAGAACGCGGAGTGTGCTGCGGCGCAAGGCGCGCG 3240  
Qy 3194 TGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3253  
Db 3241 CGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
Qy 3254 GGTGCTCATCTGCTCATCTCAAAATGCTCTCTGCGACCTCTGAGGAGCAGCCCAAAACT 3313  
Db 3301 GACTCGACACCGTGTCACTAGTGCCTCTCTGCGGTCACTCAGGACAGCCAGAGCGCA 3360  
Qy 3314 GCTGTGCGGAGCTCCCGAGGCGCAATGACCATCTTAAAGCTGAGCTGAGCTGAGCTGAGCTG 3373  
Db 3361 GCTGAGTCCGAGCTCCCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
Qy 3374 CTTAAGCAGACACTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCGCTCAGACTTCAGACCATCTCTGACTGA 3454

## RESULT 12

US-08-851-843A-173  
Sequence 173, Application US/08851843A  
Patent No. 6093809

GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6093809el Telomerase

NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 173:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4029 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: LOCATION: 1..4029  
OTHER INFORMATION: /note= "preliminary sequence for  
human TRT cDNA insert of  
OTHER INFORMATION: plasmid pGRN121"  
-08-851-843A-173  
Query Match 43.5%; Score 1520.2; DB 3; Length 4029;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;  
23 CCGGGCTTGTAGACATGACCGCGCTCTCGTGGCCCGCGGTGGCTCTCGTGGCG 82  
40 CCGGGCCACCCCGCGATCGCGCGCTCCCCGCTGCGAGCGGTGGCTCCCCGTGCG 99  
83 CAGCCGATACCGGGAGGTGGCGCGTGGCAACCTTTGTGGCGCGCTGGGGCCCCGAGG 142  
100 CAGCCACTACCGGAGTGTGCGGTGGCCACGTTCTGTGGCGCGCTGGGGCCCCAGG 159  
143 CAGCGGCTTGTGCAACCGGGGACCGGAGATCTACCGCATTTGGTGGCCGATGCT 202  
160 CTGGCGGCTGGTTCAGCGGGGACCGCGCGCTTTCCGCGGNTGGTGGGCCANTGCT 219  
203 AGTGTGATCATCTGCGGTTCACAGCTCCACTGCGACCTTTCTTCCACACAGGTGTC 262  
220 GGTGTGGTGGCTTGGGANGNANGNGGCCCCCGCGCCCTCTTCCGCGAGGTGTC 279  
263 ATCCCTGAAGAGCTGGTGGCCAGGTTGTGAGAGATCTTGGAGCGCCACGAGAGAA 322

Db 280 CTGCTCANGANGCTGGTGGCCGAGTGTCTGANGCTGTGANGCGCGCGCAANA 339  
Qy 323 CGTCTCGCTTTTGGCTTGGAGTGTCTTAACGAGGCCAGAGGGGCTTCCCATGCGCTT 382  
Db 340 CGTCTGGCTTGGCTTGGAGTGTCTTAACGAGGCCAGAGGGGCTTCCCATGCGCTT 399  
Qy 383 CACTAGTAGGCTGGTGGTGTGCTTGGCCCAACTGTATTGAGACCTCTGCTGTGAGTGG 442  
Db 400 CACCACGAGGTGGCGAGTGTCTTCCCAACGCGTGTACCGGACCTTGGGGGAGCGG 459  
Qy 443 TGCATGATGCTACTGTGAGCGAGTGGCGGAGAGCTGTGCTGTCTACTGTGTCACACA 502  
Db 460 GCGTGGGGCTGCTGCTGGCGCGCTGGCGGAGCAGCTGCTGTGCTGTGCTGTGCTG 519  
Qy 503 CTGTGCTCTTTATCTTCTGCTGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562  
Db 520 CTGCGGNTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579  
Qy 563 GTACCAATTTGTGCCACACGAGTATCTGGCCCTCTGTGCTGCTGCTGCTGCTGCTG 622  
Db 580 GTACCACTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 628  
Qy 623 CCGACCGTGGCGAGGATTTCTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 682  
Db 629 -----GGACCGGANGCTGTGGATCCACGCGCTGTGAAACCAAGCCAGCTC 670  
Qy 683 TCGCAGGAGCACCGGAAACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 742  
Db 671 -CGTCAGGAGCGCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 729  
Qy 743 TCTCAGGATGAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 802  
Db 730 CAGTGCCAGCGAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT 789  
Qy 803 GGAGGAGGAGCGCCACAGGAGGCTGTACCAACCCCATCAGGCAATCATGGTGTGCAAG 862  
Db 790 GGAGGAGGAGCGCGCTTGGCGAGGCTTGGCGAGGCTTGGCGAGGCTTGGCGAGG 849  
Qy 863 TCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 913  
Db 850 TGACCGTGGTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 909  
Qy 914 AGGAAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973  
Db 910 GGTGCGCTCTTGGCGAGGCTTGGCGAGGCTTGGCGAGGCTTGGCGAGGCTTGGCG 969  
Qy 974 CACATCTCTGCTGTACCAACCGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1033  
Db 970 CCCCCCATCCATCGCGGCGCACCGCTTGGCGAGGCTTGGCGAGGCTTGGCGAGG 1029  
Qy 1034 CAGCATTTCTTTTACTTCCAGGGAGATGCGCGAGGCTTAAACCCCTCATTTCTTCT 1093  
Db 1030 GAGACCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1089  
Qy 1094 CAGCA-ACCTCCAGCTTAACTTGAATGGGGCCAGGAGCTGGTGGAGATCATTTTCTG 1152  
Db 1090 CAATATATCTGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149  
Qy 1153 GCTAAGGCTTAGGACATCAGGACCTCTGCGAGGACACACCGCTTATCGCTGCTGCT 1212  
Db 1150 GTTCAGGCTTGGATGCCAGGA-TTCCCCGAGGTTGCCCGCTTGGCGCGCTTGGCG 1208  
Qy 1213 GCGAGATCGCGCTTGTTCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272  
Db 1209 GCGAANTGCGGCTTGTTCCTTCTGAGTGTGTTGGGACACCGCGAGTGGCCCCAC 1268  
Qy 1273 GACTCTCAGGTCACTTGTGAGGTTTGCAGAGCAACCAACAGGTGACAGTGCCT--- 1329  
Db 1269 TGTTCCTCAGAGCACTGCGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 1328  
Qy 1330 -----TGAACACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1346  
Db 1329 CCGGGAGAGCCCGAGGCTTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1388

1347 CACCTCATGATTTGCTCCGCTGCAGACAGTCCCTGGCAGTATATGTTTCTTGG 1406  
1389 GCGCTGTGAGCTGCTCCGCCAGACAGACAGCCCTGGCAGGTGTACGGCTTCGTGG 1448  
1407 GCTGTCTCTCAAGTGGTGTCTGCTAGTCTCTGGGGTACACAGGCACAATGAGCGCCG 1466  
1449 GCTGTGCTGCGCGGCTGTCGCCCCAGGCTCTGGGGCTCCAGGCAACGAAACGCCG 1508  
1467 TTCTTTAAGAACTTAAGAGTTCATCTCTGTTGGGAAATACGGGAAGCTATCATCTG 1526  
1509 TTCTCTAGGAACACCAAGAGTTTCATCTCTCGGGGAAGCATCCAAAGTCTCGTGCAG 1568  
1527 GAACTGATGTGGAAGATCAAGATGAGAGTTGCCACTGGCTCCGACAGACCCCGGGAAG 1586  
1569 GAGCTGAGCTGGAAGATGAGCTGCGGGAATGCTGCTGCTGCGAGGAGCCAGGGGTT 1628  
1587 GACCGTGTCCCGCTGCAGACACCGCTCTGAGGAGAGGATCTCTGCTAGTTCCTGTTTC 1646  
1629 GCGTGTGTTCCGGCCGACAGACACCGCTCTGCTGAGGAGATCTCTGCGCAAGTTCTTGCAC 1688  
1647 TGGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1706  
1689 TGGCTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1748  
1707 ACATTTCCAGAAAGACAGGCTCTCTTCTTACCCTAAGAGTGTGAGGCAAGCTGCAGAGC 1766  
1749 ACGTTTCAAAAGAACAGGCTCTTCTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
1767 ATTGGAGTCAAGCAACCTTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1826  
1809 ATTGGAATCAGACAGCACTTGAAGAGGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1868  
1827 AGGATCACCAGAGCACCTGCTAGCATGCCATCTGCAGAGTGTGAGTGTGAGTGTGAGTGTG 1886  
1869 AGGAGCATCGGGAAGCAGGCTCTTCTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1928  
1887 CCCAAGCGCTCGGCGCCATTTGAAACATGAGTTATGAGTGTGAGTGTGAGTGTGAGTGTG 1946  
1929 CCGTGAAGGCTCGGCGCCATTTGAAACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1988  
1947 AGAAGAGAGAGGCGCCAGCATTTCCACCGAGCTCTCAAGACTCTCTCAAGCATGCTCAAC 2006  
1989 AGAAGAAAGAGGCGGAGCTCTCACTCAGAGGTGAGGCACTGTTCAGCGTGTCTAAC 2048  
2007 TATGAGCGGAACAAACATCTCACTTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2066  
2049 TACGAGCGGCGCGCGCCCGGCTCTGCGGCGCTCTGCTGGGCGCTGACGATATC 2108  
2067 TACAGGACTCGGCGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2126  
2109 CACAGGCGCTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2168  
2127 TACTTTGTTAAGGAGATGTGACCGGGCCCTATGATGCCATCCCGAGGGTAACTGGTG 2186  
2169 TACTTTGTCAAGTGGATGTGACGGCGGTACGACACCATCCCGCAGGACAGGCTCAG 2228  
2187 GAGGTTGTTGCAATATGATCAGGCACTCGAGAGACAGTACTGTATCGCCAGTATGCA 2246  
2229 GAGGTCATCGCCAGCATCATCA---AACCACAGAACACGTAAGTGTGCTGCTGCTGCTGCTG 2285  
2247 GTGCTCCGAGAGATAGCAAGGCAAGTCCACAACTCTTTAGGAGACAGGTCACCAAC 2306  
2286 GTGCTCCAGAGAGGCGCCCATGGGCACTCGGCAAGGCTTTCAAGAGCAACGCTCTTACC 2345  
2307 CTCTCTGACCTCCAGCCATACATGGGCGAGTTTCTTAAGCATCTGAGGATTCAGATGCC 2366  
2346 TTGACAGACCTCCAGCGGTAATACCGACAGTGTGTTGGCTCACCTGCGAG-----GANAAC 2399  
2367 AGTGCACTGAGGAATCTCGTTGTGATCGAGAGAGCATCTCTATGATGAGAGCAGCAGC 2426  
2400 AGCCGCTGAGGGATGCGCTGCTCATCGAGAGAGGCTCTCTCTGATGAGGCCAGCAT 2459

Qy 2427 AGCTCTTTGACCTTCTTCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2486  
Db 2460 GGCCTCTTTCGAGCTTCTTCTACGCTTCTGCTCCACCACGCGCTGCGCATCAGGGGCAAG 2519  
Qy 2487 TGTATACGAGTGCAGGGCATCCCGCAGGCTCCAGGCTATCCACCCCTGCTCTCTGCACT 2546  
Db 2520 TCTACGTCAGTGCAGGGATCCCGCAGGCTCCATCTCTCTCCACGCTGCTCTCTGCACT 2579  
Qy 2547 CTGTGTTTGGAGACATGAGAAACAGCTGTTGCTGAGTGCAGCGGATGGGTGCTT 2606  
Db 2580 CTGTGATCGGACATGAGAAACAGCTGTTTGGGGATTCGCGGGGACGGGCTGCTC 2639  
Qy 2607 TTAGCTTTTGTGATGACCTTCTGTTGTTGAGCGCTCACTTGGACCAAGCAAAACCTTC 2666  
Db 2640 CTGGTTTGGTGGATGTTCTTGTGTTGAGCACTCACTCACTCACTCACTCACTCACTCACT 2699  
Qy 2667 CTAGACCCCTGCTCATGCGCTTCTGAGTATGGTGTGATGATGATGATGATGATGATGATG 2726  
Db 2700 CTAGACCCCTGCTCATGCGTGTCTGAGTATGGTGTGAGTATGGTGTGAGTATGGTGTG 2759  
Qy 2727 GTGTTGAACTTCCCTGCTGAGCTGCTACCTCCCTGGGTGCTGAGCTCCATACAGCTGCT 2786  
Db 2760 GTGTTGAACTTCCCTGCTGAGACAGGCGCTTGGTGTGACCGGCTTTGTTGAGATGCG 2819  
Qy 2787 GCTCACTGCTGTTTCCCTGCTGAGCTTCTGCTGCTGAGCACTCAAGCTTTGGAGGTTC 2846  
Db 2820 GCGCAGGCGCTATTCCTCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2879  
Qy 2847 TGTGACTACTCAGTTATGCCAGACCTCAATTAAGACAGGCTCACTTCCAGAGTCTC 2906  
Db 2880 AGGACTACTCTCAGTATGCCGAGCTTCCATCAGACGAGTCTCACCTTCAACCGCGC 2939  
Qy 2907 TTCAAAGCTGGGAAGACCATGCGGAACAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2966  
Db 2940 TTCAAAGCTGGGAAGACCATGCGTGCAGAACTCTTTGGGCTCTTCCGCTGAGTGTCTAC 2999  
Qy 2967 GGTCTATTTCTAGACTTTCAGGTGAGCAAGCTTCCAGAGCTCAGAGTCTGATCAATATAC 3026  
Db 3000 AGCTGTTTCTGATTTGAGGTGAGCAAGCTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTG 3059  
Qy 3027 ATCTTCTGCTTCCAGGCTTACAGTTCATGATGATGATGATGATGATGATGATGATGATG 3086  
Db 3060 ATCTTCTGCTGAGGCTTACAGTTCATGATGATGATGATGATGATGATGATGATGATG 3119  
Qy 3087 GGTGTTAGGAAGAACTCTCATCTTTCTGAGGCTATCTTCCAGCAAGCATCTCTGCTGCT 3146  
Db 3120 CAAGTTTGAAGAAACCCCACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3179  
Qy 3147 TATGCTATCTGAGGTCAAGATCCAGGAATGACACTAAAG-----GCTCTCTGCTG 3197  
Db 3180 TACTCTCTGAGAGCCAGAAAGCAGAGGATGCTGCTGGGGCCCAAGGCGCCGCGCGC 3239  
Qy 3198 TCTTTTCTCTGAGAGCCGACATTTGGTCTCTTACAGGCTTCTCTGCTCAAGTGGCT 3257  
Db 3240 CTTCTGCTTCCGAGGCGGTGAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3299  
Qy 3258 GCTCATTTCTGCTATCTACAAATGCTCTCTGAGACCTTTCAGGACAGCCCAAACTGCTG 3317  
Db 3300 CGACACGCTGCTACCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3359  
Qy 3318 TGGCGGAAGCTTCCAGAGGCGCAATGACCATCTTAAAGCTGAGCTGAGCCAGCCCTA 3377  
Db 3360 AGTCGGAAGCTTCCCGGAGCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3419  
Qy 3378 AGCAGACATTTTTCAGACCAATTTTGGACTMA 3407  
Db 3420 CCTCAGACTTCAAGACCATCTCTGACTGA 3449

RESULT 13

US-08-974-549A-292

; Sequence 292, Application US/08974549A

; Patent No. 6166178

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

## APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

## APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

## APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

## ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 292:

## SEQUENCE CHARACTERISTICS:

LENGTH: 4029 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: 1..4029

LOCATION:

OTHER INFORMATION: /note= "preliminary sequence for  
OTHER INFORMATION: human TRF cDNA insert of  
OTHER INFORMATION: plasmid pGRN121"  
US-08-974-549A-292

Query Match 43.5%; Score 1520.2; DB 3; Length 4029;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;

QY	23	CCCGGCTTGAGCAATGACCGCGCTCTCTGTTCCCGCGGGTGGCTCTCTCTGCGG	82
DB	40	CCCGGCCACCCCGCGATGCGCGCGCTCCCGCTCCGAGCGCTGCTCTCTGCGG	99
QY	83	CAGCGCATACCGGAGGTGGCGGCTTGTGCGCGGCTTGTGCGCGGCTTGTGCGG	142
DB	100	CAGCACTACCGGAGGTGGCGGCTTGTGCGCGGCTTGTGCGCGGCTTGTGCGG	159
QY	143	CAGCGGCTTGTCAACCCCGGACCGGAAGATCTACCGACTTTGGTTGGCCAAATGC	202
DB	160	CTGCGGCTGGTGACGCGCGGACCCCGCGGCTTTCGCGGCTTGTGCGGCTTGTG	219
QY	203	AGTGTGATGACACTGGGCTCACAGCTCCACCTGCGGACCTTTCCTTCCACAGGTG	262
DB	220	GGTGTGCTGCTCCCTGGGANGANGGCGGCGGCTTTCGCGGCTTGTGCGGCTTGTG	279
QY	263	ATCCCTGAAGAGCTGGTGGCGGCTTGTGCGAGACTCTGCGAGCGCAACGAGAGAA	322
DB	280	CTGCTGAANGANGCTGGTGGCGGCTTGTGCGGCTTGTGCGGCTTGTGCGGCTTGTG	339
QY	323	CGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	382
DB	340	CGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	399
QY	383	CACATAGTAGCTGCTAGCTACTTGTCCCAACATGTATTGAGACCTCGTGTCTAGTG	442
DB	400	CACACACAGCTGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	459
QY	443	TGCATGATGCTACTTGTGAGCGGAGTGGCGGAGCTGCTGCTTGTGCTTGTGCTTGTG	502
DB	460	GCGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	519
QY	503	CTGTGCTCTTTATCTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	562
DB	520	CTGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	579
QY	563	GTACCAATTTGTGCGACACCGATATCTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTG	622
DB	580	GTACCACTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	628
QY	623	CCGACCGTGGGAGGAATTTCACTAATCTTAGTCTTCTTACACAGATCAAGAGCAGTAG	682
DB	629	-----GGACCGAANGCTTGGGATCCACGGGCTTGGACCATAG	670
QY	683	TGCGCAGGAACGCAACCGCTTGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	742
DB	671	-CGTCAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	729
QY	743	TCTCACCATACAGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	802
DB	730	CAGTCCAGCGCAAGTCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	789
QY	803	GGAGGAGGAGCCACACAGGAGTGTACCAACCCCATCAGGCAAAATCATGGGTGCGCAAG	862
DB	790	GGAGGAGGAGCCCGTGGCGAGGGTCTTGGGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	849
QY	863	TCCTGCTCGGTCCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	913
DB	850	TGACCGTGTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	909
QY	914	AGGAAAGGTGTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	973
DB	910	GGGTGCTCTCTGCGACGCGGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG	969

974 CACATCTCTGCTGTCACACCCGCCAAAATGCTTTTCAGCTCAGGCCATTTATGAGAC 1033  
970 CCCCCATCCACATCGGGGCAACAGTCTTGGAACAGCCTTGTCCCGGTGTACGCC 1029  
1034 CAGACATTTCTTTACTCCAGGGGAGATGCGCAAGAGCGFTTAAACCCCTCATCTACT 1093  
1030 GAGACCAAGACATTTCTCTACTCTCAGGGGCAAGAGNACACTCGCNCCTCTCTCTACT 1089  
1094 CAGCA-ACCTCCAGCCTAACTTGAATGGGGCCAGGAGACTGGTGGAGATCACTTTCTGG 1152  
1090 CAATATATCTAGAGCCCAAGCTGACTGCGCTTGGAGGTTCTGGAGACANTCTTTCTG 1149  
1153 GCTCAAGCCTTAGACATCAGGACCACTCTCAGGACACACCGCTCTATCGCTCGATACT 1212  
1150 GTTCCAGCCTTGGATGCCAGGA-TTCCCGCAGGTTGGCCCGCTGCCCCAGCGNTACT 1208  
1213 GCGAGATGCGGCCCTGTTCCACAGCTGCTGTGAACATGCAAGATGCGCATATGTCA 1272  
1209 GGCNAATGCGGCCCTGTTCTGAGCTGCTTGGGAACCAACGCGCAGTGCCTTACGGGG 1268  
1273 GACTCCTCAGGTCACTTGCAGGTTTTCGAACAGCAAAACCAAGGTGACAGATGCT--- 1329  
1269 TGTTCCTCAAGACGCACTGCCCGCTGCGAGCTGGGTCAACCCAGCAGCGGTGTCTGTG 1328  
1330 -----TGAAACACAGCCCAACG 1346  
1329 CCGGGGAGAACCCCGAGGCTCTGTGGCGCCCGCCGAGGAGGAGGAAACACAGACCCCGT 1388  
1347 CACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGCTTTCTCGG 1406  
1389 CGCTGTGTGAGTGTCTCGCCAGCAGCAGAGCCCTTGGCAGGTGTACGGTCTGTGCG 1448  
1407 GCCTGTCTTGCAGAGTGTGTCTGTAGTCTCTGGGTACCAAGCACAATAGCGCCGC 1466  
1449 GCCTGCTGTGGCGGCTGTGTGCGCCAGGCTCTGGGCTCCAGGCACAAACGAAACGCGC 1508  
1467 TTCTTTAAGAACTTAAAGAGTTTATCTCTGTGGGAATACGCAAGTATCACTGCGAG 1526  
1509 TTCTCAGGAACACCAAGATTTATCTCCTGGGAAGATGCCAAGTCTGCTGCGAG 1568  
1527 GAATGATGGAAGATGAAGTAGAGGATTCCTACTGCTCCGACAGACCCCGGGAG 1586  
1569 GAGCTGACGTGGAAGATGAGGTGCGGAGTGGCTTGGCTGGCGAGAGCCACAGGGTT 1628  
1587 GACGCTTCCCGCTGCAGAGCAGCTCTGAGGAGAGGATCTGCTGCTACTGTTCTCTGTC 1646  
1629 GCTGTGTTCCGGCCGAGAGCAGCTGCTGCGTGGAGATCTGGCCAACTTCTGTCAC 1688  
1647 TGGCTGATGGACACATACGTGGTACAGCTGCTTAGGTCACTTTTATCATCACAGAGAGC 1706  
1689 TGGCTGATGAGTGTGTACGTCGTGAGCTGCTCAGGTCTTTCTTTTATGTACAGGAGCC 1748  
1707 ACATTCAGAGAACAGGCTCTTCTTACCTGPAAGTGTGTGAGAGAGTGTGAGAGC 1766  
1749 ACGTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTGTGAAAAGC 1808  
1767 ATTGGAGTCAGGCAACACCTTGAGAGAGTGGGCTACGGGAGCTGTCAACAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGACNTTGAAGAGGTGTGAGTCCGGAGCTGTTCGGANGCAGAGTTC 1868  
1827 AGGCATCACCAGAGACCTGGCTAGGCATGCCCATGTGCAGACTGCGTTCATCCCAAAG 1886  
1869 AGGCAGCATCGGGAAGCAGGCCCGCCCTGCTGACGTCGCAGACTCCGCTTCTATCCCAAAG 1928  
1887 CCCAAGCGCTCGGCCCATTTGTAACNTGAGTTATAGCATGGGTACAGAGCTTTGGGC 1946  
1929 CTTGACGGGTGCGGCCGATTTGTAACATGGACTAGCTGTGGAGCAGAACGTTCCGC 1988  
1947 AGAAGGAGCAGGCCAGCAATTTCAACCCAGCTCTCAAGACTCTCTTTCAGCATGTCTAAC 2006  
1989 AGAGAAAGAGGGCCGAGGCTCTCACCTCGAGGTGAAGGCACCTGTTTCAGCGTGTCTAAC 2048  
2007 TATGAGCGGACAAAACATCTCTACCTTATGGGGTCTTCTGTACTGGGTATGATGACATC 2066

2049 TAGAGCGGCGCGCGCCCGCTTCTGGGCGCTCTGTCTGGGCTGAGCATATC 2108  
2067 TACAGACCTTGGCGGCCCTTTGTGTGCTGTGCTGTGAGACACACACCCAGGATG 2126  
2109 CACAGGCGCTGGCGCACCTTGTGTGCTGTGCTGTGCGGGCCAGGACCCGCGCTGAGCTG 2168  
2127 TACTTTTAAAGCAGATGACCGCGGCTATGATGCATCCCGCCAGGTTAAGTGTG 2186  
2169 TACTTTGTCAAGTGTATGTACGGGCGGTACGACACCATCCCCCAGGACAGGCTCAG 2228  
2187 GAGGTTTGTGCCAATATGATCAGGCACCTCGGAGACAGTACTGTATCCGCCAGTATGCA 2246  
2229 GAGGTCAATGCCAGCATCATCA---AACCCAGAACACGTACTGCTGCTGCTATGCC 2285  
2247 GTGGTCCGAGAGATAGCAAGGCCAAATGCCACAGTCTTTAGGAGACAGGTACACACC 2306  
2286 GTGGTCCAGAAAGCCGCCCATGGGCAGCTCCCAAGGCTTCAAGAGCCACGTCTCTTACC 2345  
2307 CTCTCTGACTCCAGCCATACATGGCCAGTTTCTTAAAGCATCTGCAGATTTCAGATGCC 2366  
2346 TTGACAGACCTCCAGCCGTACATGCGACAGTTCTGTGGCTCACCTGCAG-----GANAAC 2399  
2367 AGTGCATCAGAGAACTCCCTTGTCTATCGAGCAGACATCTCTATGAATGAGAGCAGCAGC 2426  
2400 AGCCCGCTCAGGAGTGCCTGTCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAGCAGT 2459  
2427 AGCTGTTTGACTTCTCTGCTGACATCTCTGGGTACAGTGTGCTAAAGATTGGTGACAGG 2486  
2460 GGCTCTTTCAGAGCTTCTCTACGCTTCAATGTGCCACAGCGCTGCGCATCAGGGGCAAG 2519  
2487 TGCTATACGAGTGCAGGCAATCCCGCAGGCTCCAGCCTATCCACCTCTCTCTGAGT 2546  
2520 TCCAGCTCCAGTGCAGGGGATCCCGCAGGCTCCATCTCTCCACGCTCTCTGACAGC 2579  
2547 CTGTGTTTCGAGACATGAGAAACAGCTGTTGCTGAGGTGCAGCGGAGTGGTGTGCT 2606  
2580 CTGTGCTACGGGCAATGAGAAACAGCTGTTGGGGGATTCGGCGGGACGGGCTGCTC 2639  
2607 TTACGTTTTTGTGATGACTTCTGTGTGTGAGCGCTCACTTTGGACCAAGCAAAACCTTC 2666  
2640 CTGCGTTTTGTGTGATGATTTCTTGTGTGTGACACCTCACTCACCCAGCGCAAAACCTTC 2699  
2667 CTCAGACACCTCGTCCATGGCGTCTCTGAGTATGGTGATGATTAAGTTCAGAGACA 2726  
2700 CTCAGACACCTCGTCCGAGGTGTCCCTGAGTATGGCTGCGTGGTGAACCTTTCGGAAGACA 2759  
2727 GTGGTGAACCTTCCCTGTGAGCGCTGCTACCTCGGTGGTGTGAGCTCCATACAGCTGCT 2786  
2760 GTGGTGAACCTTCCCTGTAGAGACAGAGCCCTTGGTGGCAAGGCTTTTGTTCAGATGCCG 2819  
2787 GCTCATGCTGCTTTCCCTGTGTGTGCTGTGTGAGACATCAGACTTTTGGAGGTGTC 2846  
2820 GCCCAGGCGCTATTCCCTGTGTGCGGCTGTGTGTGATATCCCGGACCTTGGAGGTGCGAG 2879  
2847 TGTGACTACTCAGGTTATGCCAGACCTCAATTAAGCAGCGCTCACTTCCAGAGGTTC 2906  
2880 AGGCACTACTCAGCTATGCCCGGACCTTCAATCAGAGCCAGTCTCACTTCAACCCGGCG 2939  
2907 TTCAAAGCTGGGAAGACCATGCGGAACAAAGTCTCTGTGCTGCTTTGCGGTTGAAGTGTAC 2966  
2940 TTCAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTTGGGCTGAAGTGTAC 2999  
2967 GGTCTATTTCTAGACTTTCAGGTGAGACCGCTCCAGACAGTCTGCATCAATATATACAAG 3026  
3000 AGCCTGTTTCTGGATTTGAGGTGAACAGCTCCAGAGCGGTGTCACCAACATATACAAG 3059  
3027 ATCTTCTCTGCTTTCAGGCTTACAGGTTCCATGCTGTGATTCAGCTTCCCTTTGACAG 3086  
3060 ATCCCTCTGCTGACGCGTACAGGTTTTCACGCAATGTGTGTGACGCTCCCATTTTCATCAG 3119  
3087 CGTGTGTAGGAAGAACTCACATTTCTTCTGGGCTTCATCTCCAGCCAGCATCTCTGCTGC 3146

3120 CAAAGTTGGAAGACCCACACATTTTCTGGCGGTCATCTCTGACAGGGCTCCCTCTGC 3179  
3147 TATGCTATCTCAAGGTCAAGATCCAGGAATGACACTAAAG-----GCTCTTGGC 3197  
3180 TACTCATCTCAAGCCAGACAGCGAGGATGCTGGGGGCCAAGGGCGCGCGGC 3239  
3198 TCGTTTCCTCGAAGCGCACATTTGGCTCTGTACAGGCTTCTGCTCAAGCTGGCT 3257  
3240 CTTCTGCCCTCGAGCGCTGCACTGGCTGTGCCACCAAGCATTTCTGCTCAAGCTGACT 3299  
3258 GCTCATCTGCTCATCTACAAATGCTCTCTGGGACCTCTGAGGACAGCCCAAAACTGCTG 3317  
3300 CGACACCGTGTCACTACGTGCTCACTCTCTGGGTCACTCAGGACAGCCAGACGCTG 3359  
3318 TGCCGGAAGCTCCAGAGCGGACATGACATCTTTAAAGCTGAGCTGACCCAGCCCTA 3377  
3360 AGTCGGAAGCTCCCGGGACGACGCTGCTGCTGGAGCGCGGACCAACCCGCGACTG 3419  
3378 AGCAGAGCTTTCAGACCACTTTTGGACTAA 3407  
3420 CCTCAGACTTTCAGACCACTTCTGACTGA 3449

## SULT 14

-08-854-050-173

Sequence 173, Application US/08854050

Patent No. 6261836

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Haxley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 173:

SEQUENCE CHARACTERISTICS:

LENGTH: 4029 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY:

LOCATION: 1..4029

OTHER INFORMATION: /note= "preliminary sequence for

human TET cDNA insert of

OTHER INFORMATION: plasmid pORN121"

US-08-854-050-173

Query Match 43.5%; Score 1520.2; DB 3; Length 4029;  
Best Local Similarity 67.2%; Pred. No. 0;  
Matches 2317; Conservative 0; Mismatches 1028; Indels 105; Gaps 9;

QY	23	CCCGGCTTGTGACCAATGACCCCGCGCTCTCTGTCGCGCGGTGCGCTCTCTGCTGGG	82
DB	40	CCCGGCCACCCCGCGATGCGCGCGCTCCCGCTGCCGAGCGTGCCTCTCTGCTGCG	99
QY	83	CAGCCGATACCGGAGGTGTGGCGCGTGGCAACTTTGTGCGCGCGCTGGGGCCCGAGGG	142
DB	100	CAGCCTACTACCGCGAGGTGTGCGCGTGGCAGGTTCGTGCGCGCTGGGGCCCGAGGG	159
QY	143	CAGCGGCTTGTGCAACCCCGGGGACCCGAAAGATCTACCGCACTTTGGTTGCCAATGCCT	202
DB	160	CTGGCGGCTGGTGACGCGCGGGACCCCGCGGCTTTCCGCGCGTGGTGGCCCAANTGCNT	219
QY	203	AGTGTGCTGCTGCTGCGGCTCACAGCTCCACCTGCGCGACTTTCTTCCACCAGGTGC	262
DB	220	GGTGTGCTGCTGCTGCGGANGANGGCGCCCGCGCGCTCTCTTCCGCCAGGTGC	279
QY	263	ATCCCTGAAAGAGCTGTGGCGAGGTTGTGCAAGACTCTGCGAGCGCAACGAGAGAAA	322
DB	280	CTGCTCTGAANGANTGTGTGGCGCGAGTGTGCGANGCTGTGCGANGCGCGCGCAANA	339
QY	323	CGTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT	382
DB	340	CGTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTT	399
QY	383	CCTAGTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	442
DB	400	CACCAACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	459
QY	443	TGCTAGGATGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	502
DB	460	GGCGTGGGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	519
QY	503	CTGTGCTCTTTATCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	562
DB	520	CTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	579
QY	563	GTACCAAAATTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	622
DB	580	GTACCAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT	628
QY	623	CCGACCCGTGGGAGGATTTTCACTTACCTTTAGGTTCTTACCAACAGATCAAGAGCAGTAG	682
DB	629	-----GGACCCGAAANGCTGTGGATTCACCGGGCTTGGAAACCTAG	670
QY	683	TGCGCAGGAAGCACCGAAACCCCTGGCTTGCCTTCTCGAGGTACAAAGAGGATCTGAG	742
DB	671	-CGTCAGGGAGCGCGGGTCCCTTGGGCTGCCAGCCCGGGTCCGAGGAGGCGCGGGG	729
QY	743	TCTCACAGTACAGTGTGCTTCACTAAGAGGCGAGATGCTATCTCTCTCCCGAGAGT	802
DB	730	CAGTCCAGCGAAGTCTGCGGTTGCCCAAGAGCCCGAGCGGTGGCGCTGCCCTTAGCC	789

803 GGAGGGGACCCACAGGAGGTGTCTACCAACCCCATCAGGCAATCATGGTCCCAAG 862  
804 |||||  
790 GGAGCGACGCCGTTGGGAGGGGCTCTGGGCCCAACCCGGGAGGACGCCCTGGAGCCGAG 849  
805 |||||  
863 TCTGTCTCGGTCCCCCGAGGTGCTACT-----GAGAGAGAGATTTCTTTCTAA 913  
806 |||||  
850 TGACCGTGGTTTCTGTGTGTGTCACTGCGACAGCCGCGGAGAGCCACCTCTTTGGA 909  
807 |||||  
914 AGGAAAGGTCTGTGACCTGATCTCTCTGGTGGTGTGCTGTAAACACAGCCAGCTC 973  
808 |||||  
910 GGGTGCCTCTCTGGCACCGGCCACTCCACCCATCCGTGGGCGCCAGCACACCGGG 969  
809 |||||  
974 CACATCTCTGTGTACCAACCCCGCCAAAATGCTTTCAGCTCAGGCCAATTTATTGAGAC 1033  
810 |||||  
970 CCCCCATCACATCGCGGCCACACAGTCTGGGACACGCCCTTGTCCCCGGTGTACGCC 1029  
811 |||||  
1034 CAGACATTTCTTTACTTCAGGGGAGATGCCCAAGAGCGCTTAACCCCTCATTTCTACT 1093  
812 |||||  
1030 GAGACCAAGCACTTCTCTACTCTCTCAGGCGCAAGNACACTGCGNCCCTCTCTCTACT 1089  
813 |||||  
1094 CAGCA-ACCTCCAGCCTAACTTGACTGGGGCCAGGAGACTGGTGGAGATCATCTTTCTGG 1152  
814 |||||  
1090 CAATATATCTGAGGCCCAGCCTGACTGGCGTTCTGGGAGGTTCTGTGAGACANTCTTCTG 1149  
815 |||||  
1153 GCTCAAGCCTTAGACATCAAGACCACTCTGCAAGGACACACCGTCTATCGCTGATACT 1212  
816 |||||  
1150 GTTCCAGGCCCTTGGATGCCAGGA-TTCCCCGCAAGTTGGCCCGCTCTGCCCAAGCCTACT 1208  
817 |||||  
1213 GGCAGATGGGCCCTCTTCCACAGCTGCTGTGAACCATGCAGAGTGCCTAATATGTCA 1272  
818 |||||  
1209 GGCANAATGGGCCCTCTTCTGTGAGCTGCTTGGGAACCAAGCGGAGTCCCTCTAGGGG 1268  
819 |||||  
1273 GACTCTCAGGTCAATTTGAGTTTCGAACAGCAAAACCAAGAGTGCCT--- 1329  
820 |||||  
1269 TGTTCCTCAAGACGCACTGCGCGCTGGAGCTGGCGTCAACCCAGCGCGGTGTCTGTG 1328  
821 |||||  
1330 -----TGAAACCAAGCCCAAG 1346  
822 |||||  
1329 CCGGGAGAACCCCGGCTGTGTGGCGGCCCGCCGAGGAGGAGGAACACAGACCCCGT 1388  
823 |||||  
1347 CACCTCATGATTTGCTCCGCTGCAAGAGTCTCTGCGAGTATATGTTTCTTCTGG 1406  
824 |||||  
1389 CGCTGTGTGAGCTGCTCCGCCAGCACAGACGCCCTTGGCAGGTGTACGGTCTGTGCGG 1448  
825 |||||  
1407 GCCTGTCTCTCAAGGTGTGTCTGTCTGTGGGTACAGGCAATAGAGCGCGC 1466  
826 |||||  
1449 GCCTGCTTGGCCCGGCTGGTGCOCACAGGCTCTGGGGCTCCAGGCAACAGAACCGCGC 1508  
827 |||||  
1467 TTCTTTAAGAACTTAAAGAACTTCACTCTCTGGGAAATACGGCAAGCTATCACTGCAG 1526  
828 |||||  
1509 TTCTCTCAGGAACACCAAGATTTCACTCTCTGGGNAAGCATGCCAAGTCTGTGTGAG 1568  
829 |||||  
1527 GAATGATGTGGAAGATGAAGTAGAGATTTGCCATTTGGTCCGCAAGACCCCGGGAAG 1586  
830 |||||  
1569 GAGCTGACGTGGAAGATGAGGTGTGGGACTGGCGTTGGTGGCGAGGAGCCAGGGGT 1628  
831 |||||  
1587 GACCGTGTCCCGCTGCAGAGCACCGTCTCAGGAGAGAGATCTGTGCTAGTCTCTGTTC 1646  
832 |||||  
1629 GGCTGTGTTCGGCCGCAAGACACCGTCTGCTGTAGAGATCTGTGCAAGTCTCTGCAC 1688  
833 |||||  
1647 TGGCTGATGGACACATACGTGGTACAGCTCTTAGGTCACTTTTATCATCAGAGAGC 1706  
834 |||||  
1689 TGGCTGATGATGTGATGCTGTGAGCTGCTCAGGTCTTTTCTTTTATGTACGGAGACC 1748  
835 |||||  
1707 ACATTCAGAGAACAGCTCTTCTCTACCTGTAAGGTGTGTGGAGCAAGCTGCAGAGC 1766  
836 |||||  
1749 ACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
837 |||||  
1767 ATTGGAGTTCAGGCAACCTTTGAGAGAGTTCGGCTACGGAGGTGTGTCAAGAGAGGTC 1826  
838 |||||  
1809 ATTGGAATCAGACAGCACTTGAAGAGGAGTTCAGTTCGGGAGCTGTTCGGAAGCAGAGGTC 1868

QY 1827 AGGCATCACGAGACACCTGGCTAGCCATGCCATCTGCAGACTGCGCTTCACTCCCAAG 1886  
DB |||||  
QY 1869 AGGCACATTCGGGAAGCCAGGCCCGCTTGTGAGTCCAGACTCCGCTTCACTCCCAAG 1928  
DB |||||  
QY 1887 CCAACGGGCTCGGCCCATTTGTGAACATGATTTATAGCATGGGTACAGAGCTTTGGGC 1946  
DB |||||  
QY 1929 CTTGACGGGCTCGGCCGATTTGTGAACATGACTACGTCTGTGGAGCCAGAACCTTCGC 1988  
DB |||||  
QY 1947 AGNAGAGACGAGCCACCATTTTCCACAGCGTCTCAAGACTCTCTTTCAGCATGCTCAAC 2006  
DB |||||  
QY 1989 AGAAGAGAGGCCGAGCGTCTCACCTCAGGGTGAAGGCACCTTTTTCAGCGTGTCAAC 2048  
DB |||||  
QY 2007 TATGACGGGACAAAACATCTCTATGGGGTCTTCTGTACTGGGTATGAATGACATC 2066  
DB |||||  
QY 2049 TACGAGCGGCGCGGCCCGCTCTCTGGGCGCTCTGTCTGGGCTCTGACGATATC 2108  
DB |||||  
QY 2067 TACAGACCTGGGGGCTTGTGCTGGTGTGCTCTTGGACCAACACCCAGGATG 2126  
DB |||||  
QY 2109 CACAGGGCTGGGCGACCTTCTGTCTGGTGTGGGCGCCAGGACCCCGCCCTGAGCTG 2168  
DB |||||  
QY 2127 TACTTTGTTAAGGCAGATGTGACCGGGGCTATGATGCCATCCCCCAGGGTAAGCTGTG 2186  
DB |||||  
QY 2169 TACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCCCAGGACAGGCTCAG 2228  
DB |||||  
QY 2187 GAGGTGTTGCGAATATGATCAGGCACCTCGGAGACAGTACTGTATCCGCCAGTATGCA 2246  
DB |||||  
QY 2229 GAGGTCTATCGCCAGCATCATCA---AACCCAGAACACGTACTGCTGCTGCTGCTATGCC 2285  
DB |||||  
QY 2247 GTGGTCCGAGAGATAGCAAGGCCAAGTCCACAGTCTCTTTAGGAGACAGTCAACACC 2306  
DB |||||  
QY 2286 GTGGTCCAGAGGCCGCCCATGGGCACTGCGCAGGCTTCAAGAGCCACGTCTCTACC 2345  
DB |||||  
QY 2307 CTCTGTGACTCTCAGCCATACATGSGCCAGTTTCTTAAGCATCTGACAGATTCAGATGCC 2366  
DB |||||  
QY 2346 TTGACAGACTCCAGCGTACATGCGACAGTTCTGTGGTCTCACTGTCAG-----GANAAC 2399  
DB |||||  
QY 2367 AGTGCACTGAGGAACCTCGTTCTCATCGAGCAGACATCTCTATGAATGAGAGCAGCAGC 2426  
DB |||||  
QY 2400 AGCCCTCAGGATGCGCTGTCTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAGCAT 2459  
DB |||||  
QY 2427 AGCTGTTTGACTTCTCTGCACTTCTCTGCGTCAAGTGTGTGAGGTGCGAGCGGATGGTGTG 2486  
DB |||||  
QY 2460 GGCCTCTTGGACGTCTTCTACGCTTCTATGTGCCACACCGCTGCGCATCAGGGGCAAG 2519  
DB |||||  
QY 2487 TGCTATCAGAGTGCAGGCGCATCCCGCAGGCTCCAGCTTATCCACCTGCTCTGAGT 2546  
DB |||||  
QY 2520 TCTCAGTCCAGTGCAGGGATCCCGCAGGCTCCATCTCTCCAGCTGCTCTGAGC 2579  
DB |||||  
QY 2547 CTGTGTTCCGAGACATGAGAACAAAGCTGTTTGTGAGGTGCGAGCGGATGGGTGTGCTT 2606  
DB |||||  
QY 2580 CTGTGCTACGGCGACATGGAGAACAAAGCTGTTTGGGGGATTCGGCGGACGGGCTGCTC 2639  
DB |||||  
QY 2607 TTAGCTTTTGTGATGACTTCTGTGTGAGCGCTCTCTTGGACCAAGCAAAACCTTC 2666  
DB |||||  
QY 2640 CTGCTGTTGTTGATGATTTCTTGTGTGACACCTCACCTCACCCACGCGCAAAACCTTC 2699  
DB |||||  
QY 2667 CTGAGCACCTGTCTCAGTGTCTCTGAGTATGGTGCATGATATACTTGCAGAGACA 2726  
DB |||||  
QY 2700 CTGAGACCTGCTCGAGGTGCTCTGAGTATGCTGCTGGTGGTGAACCTTGGGAGACA 2759  
DB |||||  
QY 2727 GTGGTGAATCTTCTGTGAGCTGCTGCTTCCCTGGGTGTGAGCTTCCATACAGTGTGCT 2786  
DB |||||  
QY 2760 GTGGTGAATCTTCTGTGAGACGAGGCGCTTGGGTGCGACGGCTTTTGTTCAGATGCG 2819  
DB |||||  
QY 2787 GCTCACTGCTGTTTCTGCTGTGCTTGTGCTGAGACACTCAGACTTTTCAGGTGCTC 2846  
DB |||||  
QY 2820 GCCACGGGCTATTTCCCTTGTGCGGCTCTGCTGTGATACCCGAGCCCTGGAGGTGAG 2879  
DB |||||  
QY 2847 TGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACAGGCTTCACTTCCAGAGTGT 2906  
DB |||||  
QY 2880 AGCGACTACTCAGCTATGCCCGGACCTCCATCAGAGCAGTCTCACCTTCAACCGGCGC 2939  
DB |||||  
QY 2907 TTCAAAGCTGGGAGAGACCATGCGGAAACAGCTCTCTGCTGCTTTCGCTTGAAGTGTAC 2966





629 -----GGACCCGAANGCGCTGGGATCAAAGGCGCTGGAACCATAG 670  
683 TGCACAGGAGACCGAAACCCCTGGCTTGGCATCTCGAGGTACAAAGAGGCATCTGAG 742  
671 -CGTCAGGAGACCGGGGTCCCTTGGGCTGCCAGCCCGGGTGCAGAGAGCGCGGGG 729  
743 TCTCACCACTACAAGTGTGCTTCAAGTAAAGAGCCAGATGCTATCTGTCCCGAGAGT 802  
730 CAGTGCACCGCAAGTCTGCGCTGCCCAAGAGGCCAGCGCTGGCGCTGCCCTGAGCC 799  
803 GAGAGAGGAGCCCAAGCAGCAGCGGTGTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
790 GAGAGCGACCGCGGTGGGCAAGGGTCTCTGGGCCCAACCCGGGCGAGAGCGCTGACCCGAG 849  
863 TCCTGTCTGGTCCCGGAGTGCCTACT-----GCAGAGAAAGATTGTCTTTAA 913  
850 TGACCGTGTCTGTGTGGTGTACCTTGCAGACCCCGCGAGAGCCACCTCTTTGGA 909  
914 AGGAAAGGTGTCTGACCTGAGTCTCTCTGGGTGGTGTGTGTAAACAAGCCGACTC 973  
910 GGTGTGGCTCTCTGGCAGCGCACTCCACCCATCCGTGGGCGCCAGCACCAAGCGGG 969  
974 CACATCTGTGTGTACACACCGCGCAAAATGCTTTCAGCTCAGGCCATTTATGAGAC 1033  
970 CCCCCCATCCATCCGCGCCACACGCTCTGGGACAGCGCTTGTCCCGCGGTATACGCC 1029  
1034 CAGACATTTCTTTACTCCAGGGGAGATGGCCAGAGGGTCTAAACCCCTCATTTCTACT 1093  
1030 GAGACCAAGCACTTCTCTACTCTCCTCAGGCGACAAGNACACTTGGNCCCTCTTCTACT 1089  
1094 CAGCA-ACCTCAGCCTAACTTGAATGGGGCCAGGAGACTGGTGAGATCATCTTTCTGG 1152  
1090 CAATATATCTGAGGCCCAAGCTGACTGGGCTTTCGGGAGGTTTCTGGAGACANTCTTCTG 1149  
1153 GCTCAAGGCTAGGACATCAGACCACTCTGAGGACACACCGTCTATCGCGTGGTACT 1212  
1150 GTTTCAGGCTTGGATGCCAGGA-TTCCCGGAGGTGGCCCGCTGCCCGAGGNTACT 1208  
1213 GGAGATGGGCGCCCTGTTCGAAGCTCTGGTGAACCATGAGAGTGCACATATGTCA 1272  
1209 GGCAAATGGGCGCCCTGTTCGAGCTCTTGGGAAACCAAGCGAGTGGCCCTACGGGG 1268  
1273 GACTCTCAGGTACATTCAGGTTTCGAACAGCAAAACCAAGGTGACATGCCT--- 1329  
1269 TGTTCCTCAAGACGCACTCCCGCTCGAGCTGCGGTACCCAGAGCCGCTGTCTGTG 1328  
1330 -----TGAAACACAGCCCAACCG 1346  
1329 CCGGGAGAAAGCCAGGCTCTGTGGGCGCCCGGAGGAGGAGAAACACAGACCCCGT 1388  
1347 CACTCATGATTTGCTCGCTGACAGAGTCCCTGGCAGGTATATGTTTCTTCGG 1406  
1389 CGCTGGTGAGTGTCTCGCCAGACAGACGACCCCTGGCAGGTGTACGGCTTCGTGCGG 1448  
1407 GCCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACCAAGCAAAATCAGCGCGCG 1466  
1449 GCCTGCTGCGCGGTGTGTGCCCCAGGCTCTGGGCTCCAGGCAACAAGACGCGC 1508  
1467 TCTTTAAGAACTTAAGAGTTTCTGTGTGGGAAATACGGCAAGCTATCATCTGAG 1526  
1509 TTCTCAGGAACCAAGAGTTTCTCTCCCTGGGGAAGCATGCCCAAGCTCTCGTGCAG 1568  
1527 GAATGATGTGAAGATGAAGTGAAGGATTCCTACTGCTCCGAGAGCCCGGGGAAG 1586  
1569 GAGTGAAGTGAAGATGAGGTGCGGACTGCGCTTGGCTGCGAGAGCCCAAGGGTT 1628  
1587 GACCGTTCCTCCCTCAGAGCAACCGTCTGAGGAGAGATCCTGGCTACGTTCTGTTC 1646  
1629 GGCTGTGTTCCGCGCCAGAGACCGCTCTGCGTGAGGAGATCCTGGCCAAAGTCTCTGCAC 1688  
1647 TGGCTGATGACACATACCTGTGACGCTTGTAGTCAATCTTTATCATCAGAGAGC 1706  
1689 TGGCTGATGATGTGACGTGTCGAGCTGCTCAGTCTTCTTTTATGTACGAGACC 1748

1707 ACATTCAGAGAGACAGCTCTTCTTCTACCGTAAGAGTGTGTGGAGCAAGCTGCAGAGC 1766  
1749 AGTTCCTAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAGC 1808  
1767 ATTGGAGTCAAGCAACCTTGAAGAGTGGGCTACGGAGTGTGTCAAGAGAGGTC 1826  
1809 ATTGGAATCAGACAGCACTTGAAGAGGTGAGTGGGAGCTGTGGAAGCAGAGGTC 1868  
1827 AGGATCACACAGGACACTGCTAGCATGCCCATCTGCAGACTGCGCTTCATCCCAAG 1886  
1869 AGGAGCATCGGGAAGCAGGCCCGCTCTGCTGAGCTCAGACTCCGCTTCATCCCAAG 1928  
1887 CCCAACGSCCTCGCGGCCATTTGTGAACATGAGTTATAGCATGGGTACCAAGAGTTTGGC 1946  
1929 CCTGAGCGCTCGCGCGATTGTGAACATGGAATCGTCTGGGAGCCAGAACGTTCCCG 1988  
1947 AGAGGAAGAGGCGGCGAGCTTTCACCCAGGCTCTCAAGACTCTCTTCAGCATGCTCAAC 2006  
1989 AGAGAAAGAGGCGGAGCGCTCACCTCAGGGTGAAGGCACTGTTTCAGCTGCTCAAC 2048  
2007 TATGAGCGGACAAAACATCTCACTTATGGGGTCTTCTGCTACTGGGTATGAATGATC 2066  
2049 TACGAGCGGGCGCGCGCCCGCTCTCTGGCGCTCTGTGTGGCTTGGAGCATATC 2108  
2067 TACAGGACTCGCGGCGCTTGTGCTGCTGTGTGGTGTCTGTGACCAAGACCCAGGATG 2126  
2109 CACAGGCTCGCGCACTTCTGCTGCTGTGTGGGCGCCAGGACCGCGCTGAGCTG 2168  
2127 TACTTTTAAAGGAGATGTGACCGGGCTTATGATGCCATCCCGAGGTAAGCTGGTG 2186  
2169 TACTTTGTCAAGTGTGATGTGACGGGCGGTACGACACATCCCGCAGGACAGCTCACG 2228  
2187 GAGGTGTGTGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATTCGCCAGTATGCA 2246  
2229 GAGGTATCGCCAGCATCATCA---AACCAGCAACACGTACTGCTGCTGCTGATGCC 2285  
2247 GTGGTCCGAGAGTAGCAAGGCCAAGTCCACAGTCCCTTAGGAGACAGGTCACACC 2306  
2286 GTGGTCCAGAAAGCGCCCATGGGCAAGTCCGCAAGGCTTCAAGAGCCAGCTCTTACC 2345  
2307 CTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAGCATCTGCAAGATTCAGATGCC 2366  
2346 TTGACAGACCTCCAGCCGTACATCGACAGTTCGTGCTCACCTGCAG-----GANAAC 2399  
2367 ATGCACTGAGAACTCCGTTGTATGAGCAGAGCATCTATGATGAGAGCAGCAGC 2426  
2400 AGCCGCTGAGGAGTCCCGTCTCATCGAGCAGAGCTCCCTCCCTGATGAGGCGCAGT 2459  
2427 AGCCTGTTTGAATCTTCTCTGCACTTCTGCTGCTACAGTGTCTGTAAGATTGGTGACAG 2486  
2460 GGCCTCTTCGAGCTTCTTACGCTTCTATGTGCCACCAAGCCGCTGCGCATCAGGGCAAG 2519  
2487 TGCATACGCAAGTCCAGGCAATCCCGAGGCTTCAGCCTATCCACCTGCTCTGAGT 2546  
2520 TCCATAGTCCAGTCCAGGAGTCCCGAGGCTCCATCTCTCCACGCTGCTCTGAGC 2579  
2547 CTGTCTTCGAGGACATGAGCAACAAGCTGTTGCTGAGTGCAGCGGATGGTTCCT 2606  
2580 CTGTCTACGCGCATGAGAAACAAGCTGTTTCGGGGAATTCGCGGAGCGGCTGCTC 2639  
2607 TTACCTTTTGTGATGACTTCTCTGTGTGACGCTCACTTGGACCAAGCAAAACCTTC 2666  
2640 CTGCGTTTGGTGGATGATTTCTTGTGTGACACTCACCTCACCCACCGGAAACCTTC 2699  
2667 CTCAGACCTGGTCCATGGGTTCTTGTGATGGGTGATGATTAACCTGCAAGAGACA 2726  
2700 CTCAGAGCCCTGGTCCGAGGTCTCCCTGAGTATGGTGTGCTGTGAACTTCCGGAAGACA 2759  
2727 GTGGTGAACCTTCCCTGTGGAGCTTGTGAGTACCTTGGTGGTGCAGCTCCATACCAAGCTGCCT 2786  
2760 GTGGTGAACCTTCCCTGTGAGAGACGAGGCGCTTGGTGGCAGCGCTTTTGTTCAGATGCCG 2819



GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: February 2, 2004, 01:51:02 ; Search time 1398 seconds  
 (without alignments)  
 9113.998 Million cell updates/sec

le: US-09-042-460-1  
 fact score: 3496  
 uence: 1 GAATTCCGGTGGAGGCC.....CCGAGTCGGTACCAAGCTT 3496

ring table: IDENTITY NUC  
 Gapop 10.0 , Gapext 1.0

ched: 2434939 seqs, 1822278265 residues  
 al number of hits satisfying chosen parameters: 4869878

imum DB seq length: 0  
 imum DB seq length: 2000000000

t-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

abase : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1600.8	45.8	4015	9	US-09-733-294A-3
2	1600.8	45.8	4015	10	US-09-990-080-1
3	1600.8	45.8	4015	10	US-09-843-676-224
4	1600.8	45.8	4015	10	US-09-953-052-1
5	1600.8	45.8	4015	15	US-10-053-758-224
6	1600.8	45.8	4015	15	US-10-208-243-1
7	1600.8	45.8	4015	15	US-10-054-295-224
8	1600.8	45.8	4015	15	US-10-054-611-224
9	1600.8	45.8	4015	15	US-10-105-963-1
10	1600.8	45.8	4015	15	US-10-044-692-1
11	1600.8	45.8	4015	15	US-10-044-539-1
12	1600.8	45.8	4027	12	US-10-385-882-1
13	1594.6	45.6	3396	10	US-09-749-728B-32
14	1593.6	45.6	3453	10	US-10-205-629-1
15	1593.6	45.6	13766	13	US-10-105-616-1

c	16	1584.8	45.3	8742	13	US-10-105-616-6	Sequence 6, Appli
	17	1520.2	43.5	4029	10	US-09-843-676-173	Sequence 173, App
	18	1520.2	43.5	4029	11	US-09-438-486-173	Sequence 173, App
	19	1520.2	43.5	4029	15	US-10-053-758-173	Sequence 173, App
	20	1520.2	43.5	4029	15	US-10-054-295-173	Sequence 173, App
	21	1520.2	43.5	4029	15	US-10-054-611-173	Sequence 173, App
	22	1323.2	37.8	3855	15	US-10-044-692-4	Sequence 4, Appli
	23	1323.2	37.8	3855	15	US-10-044-539-4	Sequence 4, Appli
	24	1022	29.2	1856	15	US-10-294-778-11	Sequence 11, Appli
	25	724.4	20.7	1314	15	US-10-294-778-9	Sequence 9, Appli
	26	716	20.5	2176	15	US-10-044-692-3	Sequence 3, Appli
	27	716	20.5	2176	15	US-10-044-539-3	Sequence 3, Appli
	28	710	20.3	1311	15	US-10-294-778-1	Sequence 1, Appli
	29	689.8	19.7	2171	10	US-09-843-676-100	Sequence 100, App
	30	689.8	19.7	2171	10	US-09-766-253-100	Sequence 100, App
	31	689.8	19.7	2171	11	US-09-438-486-100	Sequence 100, App
	32	689.8	19.7	2171	15	US-10-053-758-100	Sequence 100, App
	33	689.8	19.7	2171	15	US-10-054-295-100	Sequence 100, App
	34	689.8	19.7	2171	15	US-10-054-611-100	Sequence 100, App
	35	470.2	13.4	873	13	US-10-282-960-1	Sequence 1, Appli
	36	468.8	13.4	4200	15	US-10-044-692-6	Sequence 6, Appli
	37	468.8	13.4	4200	15	US-10-044-539-6	Sequence 6, Appli
	38	468.8	13.4	15418	9	US-09-783-203-1	Sequence 1, Appli
	39	468.8	13.4	15418	10	US-09-994-427A-1	Sequence 1, Appli
	40	468.8	13.4	15418	11	US-09-995-419A-1	Sequence 1, Appli
	41	468.8	13.4	15418	15	US-10-141-220-1	Sequence 1, Appli
	42	468.8	13.4	15418	15	US-10-023-969-1	Sequence 1, Appli
	43	468.8	13.4	15418	15	US-10-206-447-1	Sequence 1, Appli
	44	468.8	13.4	51552	9	US-09-733-294A-30	Sequence 30, Appli
	45	338.2	9.7	519	13	US-10-282-960-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-09-733-294A-3  
 ; Sequence 3, Application US/09733294A  
 ; Patent No. US20020045588A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brett P. Monia  
 ; APPLICANT: William Gaarde  
 ; APPLICANT: Susan M. Freier  
 ; APPLICANT: Edward V. Wanciewicz  
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION  
 ; FILE REFERENCE: ISPH-0527  
 ; CURRENT APPLICATION NUMBER: US/09/733,294A  
 ; CURRENT FILING DATE: 2000-12-07  
 ; PRIOR FILING DATE: 2000-05-16  
 ; PRIOR FILING DATE: 2000-05-16  
 ; NUMBER OF SEQ ID NOS: 108  
 ; SEQ ID NO 3  
 ; LENGTH: 4015  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (56)...(3454)  
 ; US-09-733-294A-3

Query Match	45.8%	Score	1600.8	DB	9	Length	4015
Best Local Similarity	68.9%	Pred.	No. 0				
Matches	2379	Conservative	0	Mismatches	967	Indels	108
Gaps	9						
Qy	23	CCCGCGCTTGGACCAATGACCGCGCTCTCGTGGCCCGCGGTGCGCTCTCTCTCGG	82				
Db	40	CCCGCGCACCCCGCGATGCGCGCTCCCGTCCGAGCGCGTCCCTCTCTCTCGG	99				
Qy	83	CAGCGATACCGGAGGTGGCGGTGCAACCTTTGTGGCGCGCTGGGGCCCGAGGG	142				
Db	100	CAGCCACTACCGGAGGTGCTGCGCTGGCCACGTTCTGTGGCGCGCTGGGGCCCGAGGG	159				
Qy	143	CAGCGCGCTTGTGCAACCGCGGAGCCGGAAGATACCGCACTTGTGTGGTTCGCCAATGCT	202				

160 CTGGCGCTGCTGAGCGCGGGGACCGCGGCTTTCGCGCGCTGCTGCGCCAGTGCCT 219  
203 AGTGTGATGCACTGGGGCTCAGACGCTGCACCTGCGCACTTTCCTTCCACAGGTTGC 262  
220 GGTGTGCTGCTTGGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 279  
263 ATCCCTGAAAGAGCTGGTGGCGCAGGCTGTGTCAGAGACTCTGCGAGGCGCAACAGAGAAA 322  
280 CTGCTGAGAGAGCTGGTGGCGCGGAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGCGGAA 339  
323 CGTGTGCTTTTGGCTTGGCTTGAAGCTTTAAAGAGCCAGAGCGCGCGCTCCCATGCGCTT 382  
340 CGTGTGCGCTTTCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 399  
383 CACTAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442  
400 CACCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459  
443 TGCATGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 502  
460 GCGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519  
503 CTGTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562  
520 CTGCGGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 579  
563 GTACCAAAATTTGTGCCACACCGGATATGCGGCTTCTGTGCTGCTGCTGCTGCTGCTG 622  
580 GTACAGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639  
623 CGACCCGCTGGGAGGAATTTCACTAACTTAGTCTTTTCAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672  
683 TGCCGAGGAGACCGAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 742  
673 GTCAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732  
743 TCTCACAGTACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802  
733 CAGTGCAGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 792  
803 GAGGAGGAGACCCACAGAGAGTGTACCAACCCATCAGCAATATGAGTGGTGCACAG 862  
793 GAGGAGGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 852  
863 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913  
853 TGACCGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
914 AGGAAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 970  
913 GGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972  
971 CTCACAT---CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1027  
973 CCCCCATCAGATCGGGCCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032  
1028 TGAGACAGACATTTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1087  
1033 CGAGACCAAGCACTTCT 1089  
1088 CCTACTCAGCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1147  
1090 CCTACTCAGCTCTCTGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149  
1148 TCTGGGCTCAAGGCTTGAACATCAGGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1207  
1150 TCTGGGTTCCAGGCGCTGATGCCAGGACTCCCGCAGGTTGCCCGCTGCTGCTGCTGCTG 1209  
1208 ATACTGAGATGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267

1210 CTACTGGCAATGCGCGCCCTGTTTCTGGAGCTGCTTGGAAACACGCGCAGTGCCTTA 1269  
1268 TGTGAGCTCTCAGGTCACATTCAGGTTTCCAAAGCAAAACCAAGTGCACAGATGC 1327  
1270 CGGGGTGCTCTCTAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAACACCAAGCCC 1342  
1330 CTGTGCGCGGAGAACCCCGAGGCTCTGTGGCGGCCCCCGAGGAGGAGGACACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCGCGCTGACAGAGTCCCTGCGCAGGTATATGCTTTTCT 1402  
1390 CCGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1449  
1403 TCGGCGCTGCTCTCTGCAAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1462  
1450 GCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1509  
1463 CCGCTTTCTTAAAGACTTAAAGAGTTTCTGTTGGGAAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGGAACCAAGAGTTTCTTCTGCTGGGAAAGCATGCCAAGCTCTGCT 1569  
1523 GCAGAACTGATGTGGAAGATGAAAGTAGAGATTGCACTGCTGCTGCTGCTGCTGCTGCTG 1582  
1570 GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1629  
1583 GAAGACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1642  
1630 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1689  
1643 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1702  
1690 GCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749  
1703 GAGCACATTCAGAAAGAACAGGCTCTTCTTCTACCGTAAAGAGTGTGTGGAGCAAGTGCA 1762  
1750 GACCACGTTTCAAAAGAACAGGCTCTTCTTCTACCGTAAAGAGTGTGTGGAGCAAGTGCA 1809  
1763 GAGCATTTGGAGTACAGCAACACTTGAAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822  
1810 AAGCATTTGAATACAGACAGCACTTGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869  
1823 GGTGAGGATCACAGAGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1882  
1870 GGTGAGGAGCATCGGAGAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1929  
1883 CAAGCCCAACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1942  
1930 CAAGCTGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1989  
1943 GGGCAGAGGAGCAGCGCGCGCGCTTACCCAGCGTCTCAAGACTCTTTCAGCATGCT 2002  
1990 CCGCAGAAAGAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2049  
2003 CAACTATGAGCGGCAAAACATCTCCTTATGGGCTTCTTCTGCTGCTGCTGCTGCTGCTG 2062  
2050 CAACTACAGCGGCGCGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2109  
2063 CATCTACAGGACCTGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2122  
2110 TATCACAGGCGCTGCGCGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2169  
2123 GATGTACTTTTGAAGGAGATGTGACCGGCGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2182  
2170 GCTGTACTTTTGAAGGAGATGTGACCGGCGCGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2229  
2183 GGTGAGGTTTGTGCAATATGATCAGGCACTCGGAGAGCACTGCTGCTGCTGCTGCTGCTGCTG 2242  
2230 CACGAGGTCTATCGCGCAGCATCA---AACCCAGAACACGCTACTGCTGCTGCTGCTGCTG 2286  
2243 TGCAGTGTTCGCGAGATAGCCAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2302  
2287 TGCGGTGCTGCGAGAGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346

2303 CACCTCTCTGACCTCCAGCATATCGGCGCATCTTAAAGCATCTCAGGATTGCA 2362  
2347 TACCTTGACAGACCTCCAGCGGTATATCGGAGATCTCGTGGCTCAGCTGAG 2400  
2363 TGCCAGTCACTGAGGAACCTCGTTGTATCGAGCAGAGCATCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATGCGTGTATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCTGTTTACATCTCTGCACTTCTCGCTGACAGTGTGTAAGATTGGTGA 2482  
2461 CAGTGGCTCTTTCAGAGCTTCTTACGCTTCAATGTGCCACACGCGTGGCATCAGGGG 2520  
2483 CAGGTGCTATACGAGTGCCAGGCAATCCCGAGGCTCCAGGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTTACGCTGAGTCCAGGAGATCCCGAGGCTCCATCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTTTCCGAGCATGAGAACTGTTTGTGAGGTGACGGGATGGGTT 2602  
2581 CAGCTGTGCTACGCGGACATGGAGAACTGTTTGTGCGGATTCGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGTTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCTGCTTGTGATGATGATTTCTTGTGTTGTTGATGACCTCACTCAACGCGGAAAC 2700  
2663 CTCTCTCAGCACCTGCTGATGCGTCTGATGATGAGTGTGATGATGATGATGATGATGAT 2722  
2701 CTCTCTCAGGACCTGCTGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
2723 GACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2782  
2761 GACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
2783 GCTGCTCACTGCTGTTTCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2842  
2821 GCGGCGCCAGCGCTATTCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2880  
2843 GTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2902  
2881 GCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
2903 TGTCTTCAAAGCTGGGAGAACCTATGCGGAAACAGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2962  
2941 CGGCTTCAAAGCTGGGAGAACCTATGCGGAAACAGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3000  
2963 TCACGCTTATTTTACAGTGTGAGTGAACAGCTTCCAGACAGTGTGATCAATATATA 3022  
3001 TCACGCTTATTTTACAGTGTGAGTGAACAGCTTCCAGACAGTGTGATCAATATATA 3060  
3023 CAAGATCTTCTGCTTCCAGGCTTACAGGTTTCCATGATGATGATGATGATGATGATGATGATGATGAT 3082  
3061 CAAGATCTTCTGCTTCCAGGCTTACAGGTTTCCATGATGATGATGATGATGATGATGATGATGATGAT 3120  
3083 CCAGGCTTATGAGGAACCTCACTTCTTCTGAGCATCATCTCCAGGCAAGCATCTCTG 3142  
3121 TCAGCAAGTGTGAGGAACCTCACTTCTTCTGAGCATCATCTCCAGGCAAGCATCTCTG 3180  
3143 CTGCTATGCTATCTGAGTGTCAAGATCAGGATCAGGATGATGATGATGATGATGATGATGATGATGATGAT 3193  
3181 CTGCTATCTTCACTTCAAGGATCAGGATCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
3194 TGGCTCTCTTCTCTTCAAGGATCAGGATCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 3253  
3241 CGGCTCTCTGCTTCCAGGCTTACAGGTTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
3254 GGCTGCTATCTGATCTTCAAAATGTTCTTCTGAGCATCTTGAAGACAGGCAAGGCAAGGCAAGGCAAG 3313  
3301 GACTGCAACCGTGTCACTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360  
3314 GCTGTGCGGAGGATCTCCAGAGGCGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3373  
3361 GCTGAGTGGAGGATCTCCGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420

Qy 3374 CCTAAGCAGACAGCTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAGACCATCTGGACTGA 3454

RESULT 2  
US-09-990-080-1  
; Sequence 1, Application US/09990080  
; Patent No. US20020102686A1  
; GENERAL INFORMATION:  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants  
; FILE REFERENCE: 018/2586  
; CURRENT APPLICATION NUMBER: US/09/990,080  
; CURRENT FILING DATE: 1998-08-03  
; PRIOR APPLICATION NUMBER: US 09/052,864  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT) cDNA  
US-09-990-080-1

Query Match 45.8%; Score 1600.8; DB 10; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

Qy 23 CCAGGCTTGTGAGCACAATGACCGCGCTCTCGTTGCCCCCGGGTGGCTCTCTGCTCGG 82  
Db 40 CCAGGCAACCCCGCGATGCCGCGCTCCCGCTGCGAGCGCTGCGCTCTCTGCTCGG 99  
Qy 83 CAGCCGATACCGGAGGTGTGCGCGCTGCGCAACCTTTTGGGGGCGCTGGGGCCGAGGG 142  
Db 100 CAGCCACTACCGAGGTGTGCGCTGCGCAACCTTTTGGGGGCGCTGGGGCCGAGGG 159  
Qy 143 CAGCGGCTGTGCAACCCCGGAGCGCGGAGTCTACCGCATTTGTTGGTGGCCCATGCT 202  
Db 160 CTGCGGCTGTGCAACCCCGGAGCGCGGAGTCTTCCGCGCTTGGTGGCCCATGCT 219  
Qy 203 AGTGTGATGATGCTGCGGCTCAGCGCTTCCAGCTTCCAGCTTCTTCCAGCGGTGC 262  
Db 220 GGTGTGCTGCTTGGGAGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
Qy 263 ATCCCTGAAGAGCTGTGCGCGAGGTTGTGAGAGTCTTGGAGCGCGCAACGAGAGAA 322  
Db 280 CTGCTGAAGAGCTGTGCGCGAGTGTGCGCGAGTGTGCGAGGCTGTGCGAGCGCGCG 339  
Qy 323 CGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 382  
Db 340 CGTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 399  
Qy 383 CACTAGTGTGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 442  
Db 400 CACCAAGCTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 459  
Qy 443 TGCACTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 502  
Db 460 GGGTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 519  
Qy 503 CTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 562  
Db 520 CTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 579  
Qy 563 GTACCAATTTGTGCAACCGGATATCTGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622  
Db 580 GTACCAAGCTGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 639

623 CCGACCCGTGGGCGAGAAATTCATCACTTAAGCTTCTTACACAGATCAAGAGCAGTAG 582  
640 GCGTCTGGGATGC-----GACCGGCGCTGGAAACCATAG 672  
683 TCGCCAGGAGACACGAAACCCCTGGCCCTTCCATCTCGAGGTACAAAGAGGATCTGAG 742  
673 CGTCAAGGAGCGCGGGTCCCGCTGGGCTGCGACGCCCGGGTGGAGAGGCGCGGGG 732  
743 TCTCACCAAGTCAAGTGTGCTTACGTAAAGGCCAGATGCTATCTCTTCCCGAGAGT 802  
733 CAGTGCACCCGAAAGTCTGCGTTCGCCCAAGAGGCCAGGCGTGGCGCTGCCCTGAGCC 792  
803 GGAGGAGGACCCACAGCGAGGTCTACCAACCCCATCAGGCAAAATCANGGTGCCAG 862  
793 GAGCGGAGCCCGTGTGGCAGGGTCTCTGGGCCAACCCGGGAGGAGCGGTGGACCGAG 852  
863 TCTGCTGGTCCCGGAGGTGCTTACT-----GCAGAGAAAGATTGTCTTAA 913  
853 TGACCGTGGTTCGTGTGTGTCACTGCCAGACCCCGCAAGAAAGCCACCTCTTTGGA 912  
914 AGGAAGGTGTCTGACCTGAGTCTCTC---TGGGTGGTGTGCTGTAAACACAGCCAG 970  
913 GGTGGCTCTCTGGCACGCGCCATCCACCCATCCGTCGTCGGCGCCGAGCACCGCGG 972  
971 CTCCACAT---CTCTGCTGTCAACACCCCGCCAAATGCCCTTTTTCAGCTCAGGCCATTAT 1027  
973 CCCCCCATCCATCGCGCCACACAGTCCCTGGGACAGCGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACGACATTTCTTTACTCAGGGGAGTGGCCAGAGGCTTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCAGGCGA---CAAGGACAGCTGGCGCCCTCTT 1089  
1088 CTTACTCAACCACTCCAGCCCTTCACTGGGCGCCAGAGACTGGTGAGATCACTT 1147  
1090 CTTACTCACTCTCTGAGGCCAGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCT 1149  
1148 TCTGGCTCAAGGCTAGACATCAAGACCACTCTGAGGACACACCGTCTATGGGTG 1207  
1150 TCTGGGTTCCAGGCGCTTGGATCCAGGAGACTCCCGCAGGTTGGCCGCGCTGCCCGAG 1209  
1208 ATACTGGCAGATGCGGCCCTCTTCAACAGCTGCTGGTGAACCATGACAGTGGCAATA 1267  
1210 CTACTGGCAATGCGGCCCTCTTTCTGAGGCTGTTGGGAAACAGCGGACAGTGGCCCTA 1269  
1268 TGTACAGCTCTCAGTCAATGACGTTTCGACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTG 1329  
1328 CT-----TGACACACGCC 1342  
1330 CTGTGCGCGGAGAACCCAGGGCTGTGTGGGCGCCCGGAGGAGGACACAGACCC 1389  
1343 ACGCACCTCATGGATTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CGTCCCTGGTGCAGTGTCTCCGCGACACAGCAGCCCTTGGCAGGTGACGGCTTCGT 1449  
1403 TGGGCTGTCTCTGCAAGTGGTGTCTGTAGTCTCTGGGGTACCGGACAAATGAGCG 1462  
1450 GCGGGCTCTGCTGCGCGGCTGGTGCCTCCAGGGCTCTGGGGCTCCAGGCAACCGAAG 1509  
1463 CGCTTCTTTAAGAACTTAAAGATTCTATCTGTTGGGAAATACGGCAAGCTATCACT 1522  
1510 CGCTTCTCAGGACACCAAGATTCTCTCTCTGGGAGGATGCCAAGCTCTGCT 1569  
1523 GCGGAATCATGTGGAAGATGAAGTGAAGATTGCCATCTGGCTCCGCGAGCGCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGGCTGCGGCACTGCGCTTGGCTGGCGAGGCCAGG 1629  
1583 GAAGGACCTGTCCCGCTGACAGCACCGTCTGAGGAGAGGATCTGGCTAGCTTCT 1642  
1630 GGTGGCTGTGTCCGCGCGACAGACACCGTCTGCTGGTGGAGATCTTGGCCAAAGTTCT 1689  
1643 GTTCTGCTGATGGACACATACGTGGTACAGCTGCTTAGTCACTTCTTTTATCATCACAGA 1702  
1690 GCACTGCTGATGAGTGTGTACGTGCTGAGGTGCTCAGGTCTTCTTTATGTACCGA 1749  
1703 GAGCACATTCAGAAAGAACAGGCTCTTCTTACCCGTAAGAGTGTGTGAGCAAGTGA 1762  
1750 GACCACCTTTCAAGAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGA 1809  
1763 GAGCATTTGGAGTCAAGGACACACCTTGAAGAGTGGCGTACCGGAGCTGTCAAGAGA 1822  
1810 AAGCATTTGGAATCAGACAGCATTTGAAGAGGTGACAGTGGCGGAGCTGTGGAAGAGA 1869  
1823 GGTCAAGCATCACAGACACCTGGGTAGCCATGCCCCATCTGCAGACTCGGCTTCAATCCC 1882  
1870 GGTCAAGCAGCATCGGGAAGCGAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCAATCCC 1929  
1883 CAAGCCAAACGCGCTCGCGCCCATTTGTGAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCCTGACGGGCTCGCGCCGATTGTGAACATGGAATAGTCTGCTGGGAGCCAGAACGTT 1989  
1943 GGGCAGAGGAGAGCAGGCCCCAGCATTTTACCCAGCGCTCTCAAGACTCTCTTCAAGCATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCCGAGGCTCTACCTCGAGGTTGAAGGCACTGTTTCAAGCTGCT 2049  
2003 CAACATATGAGCGGCAAAACATCTCTCACTTATGGGCTCTTCTGTAAGTATGAAATGA 2062  
2050 CAACATGAGCGGCGCGCGCCCTCTCTGGGCGCTCTGCTGGGCTTGGACGA 2109  
2063 CATCTACAGACCTGGCGGCGCTTGTGTCGCTGCGTGTCTGAGACGACACACCCAG 2122  
2110 TATTCACAGGCGCTGGCGCACCTTCTGTCGCTGCTGGGCGCCAGGACCCCGCGCTGA 2169  
2123 GATCTACTTTTAAAGCAGATGTGACCGGGGCTTATGATGCCATCCCCCAGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGTGATGTGACGGGCGGTACGACACCATCCCCCAGGACAGGCT 2229  
2183 GGTGAGGTTGTGCGCAATATGATCAGGCACTCGGAGAGCAGTACTGTATCGGCCAGTA 2242  
2230 CACGAGGTCTATCGCCAGCATCATCA---AACCCAGAACACAGTACTGCTGGTGGTA 2286  
2243 TGCAGTGTCTCGGAGAGATAGCAAGGCCAAAGTCCCAAGTCTTTTGGAGACAGGTCAC 2302  
2287 TGCCGTGGTCCAGAAAGCCCGCCCATGGCAGCTCCGCAAGGCCCTTCAAGAGCACGCTCTC 2346  
2303 CACCTCTCTGACCTCAGCCATACATGGGCGAGTTCCTTAAAGCATCTCGAGATTCAGA 2362  
2347 TACCTTGACAGACTCCAGCCGTACATGCGACAGTTCTGGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCATCTGAGGAATCCGTTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAAGCCCGTGAAGGATGCGCGTCTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCGTGTGACTTCTCTGCACTTCTGCTGCTCAAGTGTGTAAGATTTGGTGA 2482  
2461 CAGTGGGCTCTTTCAGCTCTTCTCAAGTCTCAAGTGTGCGCACCAAGCTTTCGGGGGATTCGGCGGAG 2520  
2483 CAGTGTCTATACGAGTCCAGGSCATCCCCAGAGGCTCCAGCCTATCCACCTGCTCTG 2542  
2521 CAGTCTCTACGTCAGTCCAGGAGGATCCCGAGGAGTCCATCTCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGGAACAGCTGTTTGTGTGAGGTCAGCGGATGGGTT 2602  
2581 CAGCTGTGTCTACGGCGACATGGAGAACAGCTGTTTTCGGGGATTCGGCGGAGCGGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGGTGAAGCTCTCTCTCTGAGACCAAGCAAC 2662  
2641 GCTCTCGGTTGTGTGATGATTCTTGTGTGTGAGCACTCTCACTCACCACCGCGAAGAAC 2700  
2663 CTTCTCAGCAGCCTGCTCCATGCGGCTTCTGAGTATGGGTGATGATATAAATTGAGAA 2722  
2701 CTTCTCAGGACCTGCTCGAGGTTCTCTGATGATGCTGGTGGTGAATTGGGAA 2760  
2723 GACAGTGGTGAACCTTCCCTGTGAGGCGCTGGTACCCCTGGGTGGTGGAGCTCCATACAGCT 2782





503 CTGTGCTTTATCTTCTGGTCCGCCAGCTGTGCTTACAGAGTGTGTGGTCTCCCT 562  
520 CTGCGGCTCTTTGTGCTGGTGGCTCCAGCTGCGCTTACCAGGTGTGCGGCGCGCT 579  
563 GTACCAATTTGTGCCACCAAGATATCGGCCCTCTGTGCTGGTGTGTACAGGCCAC 622  
580 GTACAGCTCGGGCTGCCACTCAGGCCGCGGCCCGCCACACGCTAGTGACCCCAAG 639  
623 CGACCCGCTGGGAGGAATTTCACTAACTTAGTTCCTTAACAAGATCAAGACAGTAG 682  
640 GCCTCTGGGATGC-----GAACGGGCTTGAACCATAG 672  
683 TGCCAGGAGACACGAAACCCCTGGCTTGCATCTCGAGGTACAAGAGCATCTGAG 742  
673 CGTACGAGAGCGGGTCCCTCTGGCTGCTCCAGCCCGGGTGGAGAGCGCGGGGG 732  
743 TCTCACAGTACAAGTGTGCTTCAAGTAAAGAGCCAGATGCTATCTGTCCGAGAGT 802  
733 CAGTGCACGCGAAGTCTCCCTTGCCCAAGAGGCCAGGGGTGGCGCTGCCCTGAGCC 792  
803 GGAGAGGAGACCCACAGAGAGTGTACTACCAACCCATCAGGCAATCATGGTGGCAAG 862  
793 GGAGCGGACGCGGCTTGGGACAGGGTCTTGGGCCACCCGCGGACGAGCGGTGAGCCGAG 852  
863 TCCTGCTCGGTCCCGAGGTGCTACT-----GCAGAGAAAGATTTGTCTCTAA 913  
853 TGACCGTGTCTGTGTGTGTCACTGCGACAGCCCGCGAAGAGCCACTCTTTTGA 912  
914 AGAAGAGTGTCTGACCTGAGTCTCTC---TGGTGGTGTGTGTGTAAACAAGCCAG 970  
913 GGTGGGTCTCTGGACACGCGCACTCCACCCATCCGTGGGCGCCAGCACACGCGGG 972  
971 CTCACAT---CTCTGCTCTACCAACCCCGCAAAATGCTTTTCTGCTCAGGCTCAGGCCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCAACAGTCCCTGCGACAGGCTTGTCCCGGGTACGC 1032  
1028 TGAGACAGACATTTCTTACTTCAGGGGAGATGGCCAGAGGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAAGGCA---CAAGAGAGAGTGTGGGCGCTCTT 1089  
1088 CTTACTCAGCAACCTCCAGCTAACTTGACTGGGCGCAGAGACTTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCCAGCTGACTGAGCTCGAGGCTCTGTGAGACCATCTT 1149  
1148 TCTGGGCTCAGGCTTAGACATCAGACCACTCTGACAGACACACCGTCTATCGGTG 1207  
1150 TCTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGGCAAGTGTGCCCGCTTGCAGCG 1209  
1208 ATACTGCGAGATCGGCCCTCTGTCCACAGCTGCTGTGAACCATGACAGTGGCAATA 1267  
1210 CTACTGGCAATCGGCCCTCTGTTCTGAGCTGCTTGGAAACACGCGAGTGGCCCTA 1269  
1268 TGTCAAGCTCTCAGTCACTTGCAGGTTTTCGAACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGGTGTCTCTCAAGACCACTGCGCGCTGCGAGTGTGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGACACAGCCGC 1342  
1330 CTGTGCCGAGAGACCCAGGGTCTGTGGCGGCCCGGAGGAGGAGACACAGACCC 1389  
1343 ACCGCACTCATGATTTGCTCGGCTGCACAGCAGTCCCTGGCAGGATATGTTTTCT 1402  
1390 CGGTGCTGCTGACGTCTCTCGCCAGACACAGACCCCTGGCAGGTTACGGCTTCGT 1449  
1403 TCGGCTGTCTCTCAAGTGTGTCTGATCTCTGGGGTACAGGCAATGAGCG 1462  
1450 GCGGGCTGTCTCGCGCGGTGGTGGCCCGCAGGCTCTGGGGCTCCAGGCAACGAACG 1509  
1463 CCGTCTCTTAAAGATTTCACTCTGTTGGGGAATACGGCAAGCTATCACT 1522  
1510 CCGCTTCTCAGAAACACCAAGAGTTGATCTCCCTGGGAGAGCATGCAAGCTCTCGCT 1569

QY 1523 GCAGAACTGATGTGGAAGATGAAGTAGAGATTGCCACTGGCTCGGAGCAGCCGGG 1582  
DB 1570 GCAGAGTGAAGTGAAGATGAGCGTGGGAGCTGCGCTGGCTGGCGAGAGCCAGG 1629  
QY 1583 GAAGACCGTGTCCCGCTGCGAGACACCGCTCTGAGGAGAGATCTGGCTACGTTCT 1642  
DB 1630 GTTGGCTGTGTCCGCGCGCAGACACCGCTCTGCTGAGAGATCTGGCCCAAGTTCT 1689  
QY 1643 GTTCTGGCTGATGACACATACGTGTTGAGTGTCTAGTCTCTTTTATCATCAGA 1702  
DB 1690 GCATGGGCTGATGATGTGTAAGTGTGAGCTGCTGAGCTCTCTTTTATGTACGGA 1749  
QY 1703 GAGCACAATTCAGAGAACAGGCTCTCTTCTACCGTAAGAGTGTGAGCAAGTGA 1762  
DB 1750 GACCACTTTTCAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGCAAGTTGA 1809  
QY 1763 GAGCAATTGAGATCAGGCAACACTTGTGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
DB 1810 AAGCAATTGGAATCAGACAGCACTTTGAAGAGGTTGCACTGCGGAGCTGTGGAAGCAGA 1869  
QY 1823 GGTCAAGGCATCACAGACACCTGCGCTAGCCATCTGCGAGACTGCGCTTCAATCC 1882  
DB 1870 GGTCAAGCAGCATCGGAGAGCCAGCCGCTGCTGAGCTCCAGACTCGCTTCAATCC 1929  
QY 1883 CAGCCCAACCGCTGCGGCCATTTGTAAATGATTAAGCATGGTACGAGCTTT 1942  
DB 1930 CAGGCTGACGGCTGCGGCCGATTGTGAACATGGAATAGCTGCTGGAGCCAGAACGTT 1989  
QY 1943 GGGCAGAGAGAGCAGCCAGCACTTACCCAGCTCTCAAGCTCTCTTCAGCATGCT 2002  
DB 1990 CCGCAGAGAAAGAGGCGCGAGCTCTCACTCAGGGTGAAGGCACTGTTTCAAGCTGCT 2049  
QY 2003 CAATATGAGCGGCAAAACATCTCACTTATGGGTCTTCTGTACTGGGTATGAATGA 2062  
DB 2050 CAATACAGCGGCGCGGCCCTCTGGGCGCTCTGTGTGGCGCTTGACGA 2109  
QY 2063 CATCTACAGAGCTGCGGGGCTTTGTGCTGCTGCTGCTGCTGAGACAGACCCAG 2122  
DB 2110 TATCCAGAGGCTGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2169  
QY 2123 GATGTACTTTTGAAGCAGATGTGACCGGGGCTATGATGCCATCCCGAGGTTAAGCT 2182  
DB 2170 GCTGTACTTTTGAAGTGTGATGTGACGGCGCTGACACCATCCCGCAGGAGCT 2229  
QY 2183 GGTGAGTGTGTTGCCAATATGATCAGGCACTCGGAGAGCACTGCTATCGCCAGTA 2242  
DB 2230 CAGGAGGTCTATCGCAGCATCATCA---AACCACCAACACGCTACTGCTGCTGCT 2286  
QY 2243 TGCACTGTCTCGGAGAGATAGCCAAAGCCAAAGTCCCAAGTCTTTAGGAGACAGCTAC 2302  
DB 2287 TGCGTGTGTCAGAGCGGCCCTATGGGCACTCGCAGAGCTTCAAGAGCAAGCTCTC 2346  
QY 2303 CACCTCTCTGACCTCCAGCATAATGATGGGCTGCTTAAAGCATCTGAGGATTCAGA 2362  
DB 2347 TACCTTGACAGACCTCCAGCGTACATGCGACATGCTGGCTCACTGCTGCTGCTGCT 2400  
QY 2363 TGCCAGTCACTGAGGACTCCGTTGCTATCGAGAGAGCATCTCTATGATGAGAGCAG 2422  
DB 2401 GACCAAGCCGCTGAGGATGCGCTGCTATCGAGAGAGCTCTCTCTGATGAGGCCAG 2460  
QY 2423 CAGCAGCTGTTTGAATCTCTTCTGCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2482  
DB 2461 CAGTGGCTCTTCCAGCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
QY 2483 CAGTGTCTATGAGGAGTCCAGGCTCCCGGCTCCAGGCTTCCAGCTGCTGCTGCTGCT 2542  
DB 2521 CAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580  
QY 2543 CAGTGTGTTTCCGAGACATGAGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2602  
DB 2581 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640  
QY 2603 GCTTTTACGTTTTTGTGATGACTTTTCTGTTGGTGACGCTCACTTGGACCAACAAAC 2662

2641 GCTCTGGTTTGGTGATGATTTCTTTGGTGACACCTCACTCACCCACGGGAAAC 2700

2663 CTTTCTCAGCACCTGTGTCATGGCGTTCCTGAGTATGGTGCATGATAAACTTGCAGAA 2722

2701 CTTTCTCAGGACCTGTGTCCGAGGTGTCCCTGAGTATGCTGCTGGTGTAACTTGCAGAA 2760

2723 GACAGTGGTGAATTCCTCTGTGGAGCTGATGCTACCTCGGTGGTGCAGCTCCATACCAAGCT 2782

2761 GACAGTGGTGAATTCCTCTGTAGAGACGAGGCCCTCGGGTGGCACGGCTTTTGTTCAGAT 2820

2793 GCCTGCTCACTGCGCTGTTCCCTGGTGTGCTGTCTGTGGAACATCAGACTTTGGAAGT 2842

2821 GCGCGCCACCGGCTATTCGCCCTGGTGCAGCTGTCTGTGATAGCTACCGGACCTCGGAGGT 2880

2843 GTTCGTGACTACTCAGGTTATGCCACAGACCTCAATTAAGACGAGCCCTCACTTCCAGAG 2902

2881 GCAGAGCGACTACTCCAGCTATAGCCCGGACCTCCATCAGAGCGAGTCTCACCTTCAACCG 2940

2903 TGTCTTCAAAGCTCGGGAAGACCATCGGAAACAAGCTCTCTGTCGGTCTTTCGGGTTCAGAGT 2962

2941 CGGCTTCAAGGCTGGGAGAACATGGTCGCAAACTCTTTGGGGTCTTTCGGCTGAAGTG 3000

2963 TCACGGTCTATTTCTAGACTTGCAGGTGAACAGCTTCCAGACAGTCTGCATCAATATATA 3022

3001 TCACAGCCTGTTTCTGAGTTTGCAGGTGAACAGCTTCCAGACGGTGTGCACCAACATCTA 3060

3023 CAAGATCTTCCGTCTCAGGCTTCCAGGTTCCATGCATGTGTGATTCAGTCTCCCTTTGA 3082

3061 CAAGATCTCTCTGCTCAGCGGTACAGGTTTACGCGATGTGTGCTGCAAGCTTCCCATTTCA 3120

3083 CCAGCGTGTAGGAAGAACTCACATTTCTTGGGCATCATCTCCAGCCAAAGCATCCTG 3142

3121 TCAGCAAGTTTGGAGAACCCACATTTTCTTCGGCTCATCTGACGGCTCCCT 3180

3143 CTGCTATGCTATCTGAAGTCAAGATCCAGGAATGACATAAG-----GCTC 3193

3181 CTGTACTCCATCTCTGAAGCCAAAGAACGAGGATGTCTGCTGGGGCCAAAGGCGCGC 3240

3194 TGGTCTCTTCTCTGAAGCCGCAATTTGGCTTGTCTACGAGGCTTCTGCTCAAGCT 3253

3241 GGGCCCTCTGCCCTCCAGGCGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGTCTCAAGCT 3300

3254 GGCTGCTCATTTCTGTCATCTACAAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313

3301 GACTCGACCGTGTCACTTACGTGCCATCTCTGGGGTCACTTCAGGACGCCCCAGACGA 3360

3314 GCTGTGCGGAAGCTCCAGAGCGCAATGACCACTCTTAAAGCTGCAGCTGACCCAGC 3373

3361 GCTGAGTCGGAAGCTCCGGGACGAGCTGACTGCCCTTGGAGCCGCGACCAACCCGGC 3420

3374 CCTAAGCACAGCTTTCAGACCAATTTTGAAGTAA 3407

3421 ACTGCCCTCAGACTTCAAGACCACTCTCGAAGTCA 3454

STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 14-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/052,919  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 13-NOV-1997  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "human telomerase  
transcriptase (hTRT)"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-953-052-1

Query Match	45.8%;	Score 1600.8;	DB 10;	Length 4015;
Best Local Similarity	68.9%;	Pred. No. 0;		
Matches 2379;	Conservative	0;	Mismatches 967;	Indels 108; Gaps 9;

  

QY	23	CCCGGCTTGAGCACAATGACCCGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCTGCG	82
Db	40	CCCGGCCACCCCGCATGCGCGCGCTCCCGCTGCCAGCGGTGCTCCCTGCTGCG	99
QY	83	CAGCGGATACCGGAGGTGTGGCCGCTTGCAACTTTGTGCGGCGCTTGGGCGCCCGAGGG	142
Db	100	CAGGCATACCGCGAGGTGTGCGGTGTGCACAGTTGTGTCGCGCGCTTGGGCGCCCGAGGG	159

ULT 4  
 09-953-052-1  
 sequence 1, Application US/09953052  
 atent No. US20020173476A1  
 GENERAL INFORMATION:  
 APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Hatley, Calvin B.  
 Andrews, William H.  
 TITLE OF INVENTION: Antisense Compositions for Detecting and  
 Inhibiting Telomerase Reverse Transcriptase  
 NUMBER OF SEQUENCES: 72  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: Townsend and Crew LLP  
 Townsend and Townsend and Crew LLP



2287 TCCGCTGCTCCAGAGGCGCCGCCATCGGCACGTCGCAAGGCCCTTCAAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGCGCCAGTCTCTTAAGCATCTCAGGATTCAGA 2362  
2347 TACCTTGACAGACCTTCCAGCGGTATCGACAGTTCTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGTGCTAGAGAACTCCGTTGTTCATCGACGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCAGCCCGCTGAGGATCGCGTCTCATCGACGAGCTCCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCCCTGTTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGCCTCTTTCGACGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2483 CAGGTGCTATACGAGTCCAGGCGCATCCCCAGGCGTCCAGCTATCCACCTGCTCTG 2542  
2521 CAAGTCTACGTCCAGTCCAGGCGATCCCGCAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTCCGAGACATGAGAACAGCTGTTGCTGAGTGAGCGGATGGTT 2602  
2581 CAGCTGTGCTACGCGGACATGAGAACAGCTGTTGCGGGGATTCGGCGGACGCGCT 2640  
2603 GCTTTTACGCTTGTGTGATGACTTCTGTTGTCAGCGCTCAGTTGACCAACAAAC 2662  
2641 GCTCTGCTGTTGTGATGATTTCTGTTGTCAGCTCAGCTCAGCTCAGCTCAGCT 2700  
2663 CTCTCTGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2722  
2701 CTCTCTGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2723 GACGTGTGAACTCTCTGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2782  
2761 GACGTGTGAACTCTCTGCTGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
2783 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2842  
2821 GCGGCCCCAGCGCTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
2843 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2902  
2881 GCAGAGGACTACTCCAGTATGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940  
2903 TGTCTTAAAGCTGGAGAACCATCGCGAAACAGCTCCTGCTGCTGCTGCTGCTGCTG 2962  
2941 CGGCTTCAAGCTGGAGAACATCGCTGCGAACTCTTTGGGGTCTTGGGGTGAAGTG 3000  
2963 TCAGGCTGCTATTTAGACTTGGAGTGAACAGCTCAGAGCTGCTGCTGCTGCTGCTG 3022  
3001 TCACAGCTGTTTCTGGATTTGCAAGTGAACAGCTCAGAGCTGCTGCTGCTGCTGCT 3060  
3023 CAAGATCTTCTGCTTCAAGGCTACAGGTTCCATGCTGCTGCTGCTGCTGCTGCTG 3082  
3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
3083 CAGGCTGTTAGAGAACCTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3142  
3121 TCAGCAAGTTTGGAGAACCCACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180  
3143 CTGCTATGCTATCTCTGAGGTCAGAAATCCAGGAATCAGACTAAG-----GCCTC 3193  
3181 CTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240  
3194 TGGCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3253  
3241 CCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3300  
3254 GCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3313  
3301 GACTCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3314 GCTGTGCGGAGCTCCAGAGGCGCAATGACCATCTTTAAAGCTGCTGCTGCTGCTGCT 3373  
3361 GCTGAGTGGAGCTCCGCGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420

QY 3374 CCTAAGCACAGACTTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTCCCTCAGACTTTCAGACCATCTCTGGACTGA 3454

## RESULT 5

US-10-053-758-224  
; Sequence 224, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; ; Lingner, Joachim  
; ; Nakamura, Toru  
; ; Chapman, Karen B.  
; ; Morin, Gregg B.  
; ; Harley, Calvin  
; ; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-0029300US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 56..3454  
; OTHER INFORMATION: /product= "hTRT"  
; /note= "human telomerase reverse  
; transcriptase (hTRT) catalytic protein  
; component"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-10-053-758-224

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 58.9%; Pred. No. 0;

Matches	2379;	Conservative	0;	Mismatches	967;	Indels	108;	Gaps	97;
23	CCCGGCGCTTGAGCACAATGACCCGGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCTGCG	82							
40	CCGGGCGACCCCGCGGATGCCGCGCGCTCCCGCGTGCAGACGCTGCGCTCCCTGCTGCG	99							
83	CAGCGGATACCGGGAGTGTGCGCGCTGGCAACTTTGTGTGGCGCGCTGGGGCCCGAGGG	142							
100	CAGCCACTACCGCGAGTGTGCCCTGGCACCGTCTGGCGCGCTTCCGCGCTGTGGCCAGTGGCT	159							
143	CAGCGGCTTGTGCAACCCGGGGACCCGGAAGATCTACCGCACTTTTGTGGCCCAATGCTCT	202							
160	CTGGCGCTGGTGCAGCGGGGACCCGGCGCTTCCGCGCTGTGGCCAGTGGCT	219							
203	AGTGTGCATGCACCTGGGCTCACAGCTCCACCTGCCGACCTTTCTTCCACCAAGTGTCT	262							
220	GGTGTGGTGTCCCTGGGACGCAACGCGCGCGCCCGCGCGCCCTCTCTCCGCAAGGTGTC	279							
263	ATCCCTGAAGAAGCTGTGGCGCAGGGTGTGCAGAGACTCTGCAGCGCAACGAGAGAAA	322							
280	CTGCCCTGAAGAGCTGGTGGCCCGAGTGTGCAGAGCTGTGCAGCGCGGGCGCGAAGAA	339							
323	GTGTCTGGCTTTTGGCTTTTGAAGCTCTTAAAGAGGCCAGAGCGGGCTTCCCATGGCCCTT	382							
340	CGTCTCGGCTTGTGGCTTGTGGCTGTCTGGACGGGGCCCGCGGGCGCCCGCGAGGCTT	399							
383	CACCTAGTAGCGTGTAGTACTCTTCCGCCAACACTGTTATTGAGACCTTCGCTGTCAGTGG	442							
400	CACCAACGCGTGGCAGACTACTGCCCAACACGCTGACCGACGCACTCGGGGAGCGG	459							
443	TGCATGGATGCTACTGTTGAGCCGAGTGGGCGACGACCTGCTGGTCTACTCTGTGCGACA	502							
460	GGCGTGGGGGCTGTCTGTGCGCGCGTGGGCGACGAGTGTCTGGTTCACCTGTGGCAG	519							
503	CTGTGCTCTTATCTTCTGGTGTGCCCCCAGCTGTGCTTACAGGTGTGTGGCTCTCCCT	562							
520	CTGGCGCTCTTGTGTGTGGTGTGCGCTGCGCTACCAAGTGTGGGCGCGCGCT	579							
563	GTACCAAAATTGTGCGCACCAACGATATCTGGCCCTCTGTCTCGCTAGTTTACAGGCCCCAC	622							
580	GTACCGAGCTGGCGCTGCCACTCAGGCCCGGCCCGCCGACACGCTAGTGGACCCGAG	639							
623	CCGACCGTGGGCGAGAAATTTCACTAACCTTAGGTTCTTACAACAGATCAAGAGCAGTAG	682							
640	CGCTCTGGGATGC-----GAAACGGCGCTGGAACCATAG	672							
683	TCGCCAGGAGACACGAAACCCCTGGCTCTCCCATCTCGAGGTTACAAAGAGCATCTCGAG	742							
673	CGTCAGGGAGGCGGGTCCCTCTGGGCTCTCCAGCCCGGTTGCGAGAGGCGCGGGG	732							
743	TCTCAACAGTACAGTGTGCTTACGCTTAAGAGGCGCAGATGCTATCTCTGTCCCGAGAGT	802							
733	CAGTGCAGCGCAAGTCTCCCTTCTCCCAAGAGGCCAGCGGTGGCGCTGCCCTTGAGCC	792							
803	GGAGGGAGGCCCAACAGCAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGCCAG	862							
793	GGAGCGAGCGCCCTTGGGCGAGGGTCTGTGGGCCACCCGGGCGAGACGCGTGGACCCAG	852							
863	TCCTGTCTGGTCCCCGAGGTGCCTACT-----GCAGAGAAAGATTTGTCTTCTTAA	913							
853	TGACCGTGGTTTTCTGTGTGGTGTCACTTGCAGACCGCGGAGAGACCACTCTTTGGA	912							
914	AGGAAAGTGTCTGACCTGAGTCTCTC---TGGGTGGTGTGTGTAAACACAGCCCGAG	970							
913	GGGTGTGCTCTCTTGGCAACGCGCACTCCCAACCATCCGTGGGCGCGCAGCACCGCGG	972							
971	CTCCACAT---CTCTGCTGTCAACACCCCGCAAAATGCTTTTTCAGCTCAGGCCATTTAT	1027							
973	CCCCCATCACATCGGCGCACACGTCCTCTGGGACACGCTTGTCTCCCGGTGTACGC	1032							
1028	TGAGACAGACATTTCTTTTATCTCAGGGGAGATGGCGCAAGAGCGTCTAAACCCCTCAT	1087							
1033	CGAGACCAAGCATCTCTCTACTCTCTCAGGGGA---CAAGGAGCAGCTCGCGCCCTCT	1089							

2123 GATGTACTTTTAAAGCAGATGTGACCGGGGCTATGATGCCATCCCCAGGGTAAGCT 2182  
2170 GCTGTACTTTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCCCCAGGACAGCT 2229  
2183 GGTGGAGGTGTTGGCCATATGATCAGGCACTCGGAGACAGTACTGTATCCCGCAGTA 2242  
2230 CACGGAGGTCACTCGCCAGCATATCA---AAGCCAGAAACACGTACTCGGTGCGTGA 2286  
2243 TGCAGTGGTCCGAGAGATAGCAAGCCCAAGTCCACAAAGTCTTTAGGAGACAGGTCA 2302  
2287 TGCCTGTGTCAGAGGCGCCCATGGGACGTCCGCAAGGCTTCAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATAGGCGCAGTTCCCTTAAGCATCTGCAAGATCAGA 2362  
2347 TACCTTGACAGACCTCCAGCCGTACATGCGACAGTTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCAGTGCCTGAGGAACCTCGTGTGTCAGCAGACGATCTCTATGAATGAGAGCAG 2422  
2401 GACAGCCCGCTGAGGGATGCGTGTGATGAGCAGAGCTCTCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCTTTTGAATCTTCTGCACTTCTGCGTCAAGTGTGCTAAAGATTGTTGA 2482  
2461 CAGTGGGCTCTTCCAGCTCTTCTAGCGTTCTATGTGCCACACGCGGTGCGCATCAGGG 2520  
2483 CAGTGTATACGAGTGCAGGCACTCCCGAGGCTCCAGCTTATCCACCTGCTCTG 2542  
2521 CAGTCTAGTCCAGTGCAGGGATCCCGAGGCTCCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTGTGTTCGAGACATGAGAACAGCTGTTGCTGAGTGCAGCGGATGGGTT 2602  
2581 CAGCTGTGTACGGGACATGAGAACAAAGTGTGTTGGGGATTCGCGGGAGCGGCT 2640  
2603 GCCTTTACGTTTGTGATGACTTCTGTTGTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCCTCGGTTTGGTGGATGATTCTTGTGTGTGACACCTCACTCACCACCGCGAAAC 2700  
2663 CTTCCTCAGCACCTCGTCCATGCGCTTCTGATGATGGTGCATGATAAACTTGCAGAA 2722  
2701 CTTCCTCAGACCTCGTCCGAGGTCTCCCTGATATGCTGCTGGTGAATTCGGAA 2760  
2723 GACAGTGGTGAATCTCCCTGTGAGGCTGGTACCTGGGTGTGAGTCCATACCACT 2782  
2761 GACAGTGGTGAATCTCCCTGTGAGACGAGGCTGGGTGGACAGGCTTGTGTTCAAT 2820  
2783 GCCTGTCTACTGCTGTTTCCCTGGTGGTCTGCTGAGACCTCAGACTTTGGAGGT 2842  
2821 GCGGCGCCAGGCTATTCCTTGGTGGGCTGCTGCTGGATACCCGACCTCGAGGT 2880  
2843 GTTCTGTGACTACTAGGTATGCCAGACCTCAATTAAGAGCGCTCACCCTTCCAGAG 2902  
2881 GCAGAGCGACTACTCCAGCTATGCCCGACCTCCATCAGAGCGAGTCTCACTCAACCG 2940  
2903 TGTCTTCAAGCTGGGAAGCATTGGGAACAGCTCCCTGCTGCTGCTTGGCGTGAAGTG 2962  
2941 CGGCTTCAAGCTGGGAGAAATGCTGCGAACTCTTTGGGGTCTTGGCGCTGAAGTG 3000  
2963 TCACGCTATTTCTAGACTTGCAGGTGAACAGCTCCAGACAGTCTGCATCAATATA 3022  
3001 TCACAGCTGTTTCTGATTTGAGTGAAGTGAAGCTCCAGAGGTGTGCAACAACATCTA 3060  
3023 CAAGATCTTCTGCTTCCAGGCTACAGTTCCATGATGTGTGATTAAGTTCCTTTGA 3082  
3061 CAAGATCTTCTGCTTCCAGGCTACAGTTTCAAGCTGTTGCTGCAAGCTCCCATTTCA 3120  
3083 CCAGCTGTGTTAGGAAGAACTCATTCTTCTGGGCTATCTCCAGCCCAAGCATCTG 3142  
3121 TCAGCAAGTTTGGAGAACCCACATTTTCTGCGCTCACTCTGACAGGCTCTCT 3180  
3143 CTGCTATGCTATCTTGAAGGTCAAGATCCAGGAATGACATAAAG-----GCCTC 3193  
3181 CTGCTACTCTTCTTGAAGAACCAAGACGAGGATGTCTGTTGGGGCCAAAGGCGCGC 3240  
3194 TGGCTCTTCTTCTTGAAGCGCACATTGGCTTACGAGGCTTCTGCTCAAGCT 3253

Db 3241 CGGCCCTTGCCTCCGAGGCGTGCAGTGGCTGTGCCACCAAGCATTTCTGCTCAAGCT 3300  
QY 3254 GGTGCTCTATTCTGTCACTTCAAAATGTCTCTGGAGCTCTGAGGACAGCCCAAAACT 3313  
Db 3301 GACTGCACACCGTGTCACTAGTGCCACTCTCTGGGTCACTCAGGACAGCCAGAGCA 3360  
QY 3314 GTGTGCCGAAGCTCCAGAGGCGACAAATGACCATCTTAAAGCTGAGCTGACCCAGC 3373  
Db 3361 GTGAGTGGAGAGCTCCCGGGGACGACGCTGACTGCTTGGAGCGCGACCAACCCGCG 3420  
QY 3374 CTTAAGACAGACTTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454

RESULT 6  
US-10-208-243-1  
; Sequence 1, Application US/10208243  
; Publication No. US20030044394A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaeta, Federico C.A.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune  
; TITLE OF INVENTION: Response to a Telomerase Antigen  
; FILE REFERENCE: 015389-003500PC  
; CURRENT APPLICATION NUMBER: US/10/208,243  
; CURRENT FILING DATE: 2002-07-30  
; PRIOR APPLICATION NUMBER: US/09/675,321  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/112,006  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898  
; PRIOR FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)  
US-10-208-243-1

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGGCCCTTGAGCAATGACCCGGCTCCTCGTTGCCCGCGGTGGCTCTCTGCTGG 82  
Db 40 CCGGCCCAACCCCGGATGCGCGCTCCCGCTGCCGAGCGGTGCGCTTCTGCTGG 99  
QY 83 CAGCGCATACCGGAGGTGTGGCGCTGGCAACCTTTGTGCGGCGCTGGGGCCGAGGG 142  
Db 100 CAGCCACTACCGGAGGTGTGCGCTGGCCACGTTGCTGCGGGCTGGGGCCGAGGG 159  
QY 143 CAGGGCGTGTGCAACCCGGGACCCGGAAGATCTACCGACATTTGTTGCCCAATGCT 202  
Db 160 CTGGCGGTGTGCAAGCGCGGGACCCGGCGGCTTTCGCGCGCTGGTGGCGAGTGC 219  
QY 203 AGTGTGATGCACTGGGGCTCACAGCTCTCACCTGCGGACCTTTCTTCCACCGAGTGC 262  
Db 220 GGTGTGCTGCTTGGAGACGACGCGCGCGCCCGCCCTCTCTTCCGCGAGTGTG 279  
QY 263 ATCCCTGAAAGAGTGTGTGGCCAGGGTTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTGAAAGAGTGTGTGGCCAGTGTGTGAGAGCTGTGCGAGCGCGCGCGAGAA 339  
QY 323 GTGTGCTGCTTTTCTGCTTGTGAGTGTGTTACGAGGCGCAGAGGCGCTCCCATGGCTT 382  
Db 340 GTGTGCTGCTTCTGCTTGTGAGTGTGAGGCGCGCGCGCGCGCGCTTCTTCTGAGGCTT 399



393 CACATAGTACGTCGTCAGTACTTGGCCAAACATGTTATTGAGACCCCTCGGTGTCAGTGG 442  
400 CACCACAGCGTCGTCAGTACTTGGCCAAACATGTTATTGAGACCCCTCGGTGTCAGTGG 459  
443 TGCATGATGTCATCTGTTGAGCGAGTGGCGACGACCTGCTGCTACTCTGTCGTCACA 502  
460 GCGTGGGGTGTCTGCTGCGCGCTGGCGGACGACGTCGCTGCTCACTGCTGGCAG 519  
503 CTGTGCTCTTATCTTCTGTTGTCGCCCCCAGCTGTCCTACAGGTGTGGGTCTCCCT 562  
520 CTGCGCGCTCTTGTGTGTGGTCTCCAGCTGCGCTACAGGTGTGGGGCGCGCT 579  
563 GTACCAAAATTTGTCACCAAGGATATCTGCGCTCTGCTGCTGCTAGTACAGGCCAC 622  
580 GTACCAAGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 639  
623 CCGACCCCTGGCGAGGAATTTCACTAACCTTAGGTTCTTAAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGATG-----GAACGGCGCTGGAACCATAG 672  
683 TCGCCAGGAACAACCGAAACCCCTGGCTTCCCATCTCGAGGTACAAAGGACATCTGAG 742  
673 GGTACGGAGGCGGGGTCTCCCTTGGGCTTCCAGCCCGGGTGGAGAGGCGCGGGG 732  
743 TCTCAACAGTACAAGTGTGCTTCCAGCTAAGAGGCCAGATGCTATCTCTGTCCTGAGGT 802  
733 CAGTCCAGCGAGTCTGCGTTGCGCAAGAGGCCCGAGCGTGGCGCTGCCCTGAGCC 792  
803 GGAGAGGAGGACCCACAGGAGGTGCTACCAACCCCATCAGGCAAAATCATGGTGCCAG 862  
793 GGAGCGAGCGCGTGTGGCGAGGGGTCTCGGCCCAACCCCGGCGAGACGCGTGGACCGAG 852  
863 TCCTGCTCGTCCCCCGAGGTGCTACT-----GCAGAGAAAGATTGTCTTCTAA 913  
853 TGACCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
914 AGGAAGGTGTCTGACCTGAGTCTCTC-----TGGGTGCTGCTGCTGCTGCTGCTGCTG 970  
913 GGGTGGCTCTCTGCGCAGCGCCACTCCACCCATCCGTCGGCGCGCCAGCACACCGCGG 972  
971 CTCACAT-----CTGCTGCTGCTCACCACCCCGCAAAATGCTTTTCACTCAGGCGCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCGACACAGTCCCTGGGACACGCTTGTCTCCCGGTGACGC 1032  
1028 TGAGACCAAGATTTCTCTTACTCCAGGAGATGCGCAAGCGCTTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCCAGCTTAACTTGACTGGGGCCAGGAGACTGCTGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCCAGGCTGACTTGGGCTCTGAGAGCTCTGAGAGACCATCT 1149  
1148 TCTGGCTCAAGGCTTAGGACATCAGGACCACTCTGAGGACACCGCTTATTCGGCTCG 1207  
1150 TCTGGGTTCAGGGCCCTGATGTCAGGAGCTCCCGCGAGTTCGCCCGCTGCCCGCGG 1209  
1208 ATACTGGAGATGGGGCCCTGTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1267  
1210 CTACTGGCAATATGGGGCCCTGTTCTGAGCTGCTTGGGAAACAGCGCAGTGGCCCTA 1269  
1268 TGTGAGCTCTCAGGTCAATTCAGGTTTCGAAACAGCAAAACCAAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGACACAGCCC 1342  
1330 CTGTCCCGGAGAGCCCAAGGCTGTGTGGGGCCCGCCCGAGAGGAGGACACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCCGCTGACAGCAGTCCCTGGCAGGTATATGGTTTCT 1402  
1390 CGTGGCTGTGTGAGTGTCTCCCGCAGCAGCAGCAGCCCTTGGCAGGTGTACGGCTTCT 1449

Qy 1403 TCGGGCTGTCTCTGCAAGGTGTGTCTGTCTGTCTGTGGGTACAGGCACAATGAGCG 1462  
Db 1450 GCGGGCTGTCTCTGCGCGCTGTGTCTCTCCCAAGGCTCTGTGGGCTCCAGGCAACAGAACG 1509  
Qy 1463 CCGCTTCTTTAAGAACTTAAAGAGTTTCTCTGTGTGGGAAATACGGCAAGCTTACT 1522  
Db 1510 CCGTCTCTCAGGAACACCAAGAGTTTCTCTCCCTGGGAAGCATGCCAAGCTCTCGT 1569  
Qy 1523 GAGGAATCTGATGTGGAAGATGAAGATGAGAGATTTGCCATCTGGCTCCGACAGCCCGG 1582  
Db 1570 GCAGGAGCTGACGTGGAAGATGAGGTGCGGACCTGCGCTTGGCTGCGCAGAGCCAGG 1629  
Qy 1583 GAAGGACCGTGTCCCGCTGACAGACACCGTCTGAGGAGAGGATCTGGCTTACGTTCT 1642  
Db 1630 GGTGGCTGTGTCTCGGCGCGAGAGACCGTCTGCTGAGGAGATCTTGGCCAGTTCT 1689  
Qy 1643 GTTCTGGCTGATGGAACATACGTGGTACAGCTGTCTTAGTCTTCTTTTATCATACAGA 1702  
Db 1690 GCCTGGCTGATGAGTGTGTAGCTGTCTGAGCTGTCTGAGTCTTCTTTTATGTACCGGA 1749  
Qy 1703 GAGCATACTCCAGGAACACAGGCTCTTCTTCTACGTAAGAGTGTGAGGCAAGCTGCA 1762  
Db 1750 GACACGTTTAAAGAACAGGCTCTTCTTCTACCGAAGAGTGTCTGAGCAAGTTGCA 1809  
Qy 1763 GAGCATTTGAGTCCAGCAACACCTTGAAGAGTGTGGGTACCGGAGCTGTCAACAGAGA 1822  
Db 1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGCTGAGCTGCGGAGCTGTGGAAGCAGA 1869  
Qy 1823 GGTGAGCATACACAGGACACCTGCTAGCCATGCCATCTGCACACTGCGCTTCATCCC 1882  
Db 1870 GGTGAGGAGCATCGGAAGCAGGCGCCCTGTGAGTCCAGACTCCGCTTCATCCC 1929  
Qy 1883 CAAGCCCAAGCGCTCGGCCCATTTGTGAACATGATTATAGCATGGGTACCAAGCTTT 1942  
Db 1930 CAAGCTGACGGCTCGGCCCATTTGTGAACATGACTACGCTGCTGGGAGCCAGAACGTT 1989  
Qy 1943 GGGCAGAAAGAGCAGGCGCCAGCATTTACCCAGCGTCTCAGACTCTCTTCAGCATGCT 2002  
Db 1990 CCGCAGAAAGAGGCGGAGCGTCTACCTCGAGGTTGAAGCACTGTTCAGCGTCT 2049  
Qy 2003 CAATATGAGCGGACAAAACATCTTCACTTATGGGCTTCTCTGCTGCTGGGTATGAATGA 2062  
Db 2050 CAACTAGAGCGGCGCGGCGCCCGGCTCTCTGCTGCTGGGCTTGGACGA 2109  
Qy 2063 CATCTACAGACCTGCGCGGCTTGTGTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2122  
Db 2110 TATCCACAGGCGCTGCGCACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2169  
Qy 2123 GATGTACTTTTAAAGCGAGATGTGACCGGGGCTTATGATGCCATCCCCCAGGGTAAAGCT 2182  
Db 2170 GCTGTACTTTGTCAAGGTGATGTGACGGCGCGTACGACACCATCCCCCAGGACAGGCT 2229  
Qy 2183 GGTGGAGGTTTGTCCAAATATGATCAGGCACTCGGAGAGCAGTACTGTATTCGCGCAGTA 2242  
Db 2230 CAGGAGGTCTATCCGACGATCATCA---AAACCCAGAACACGACTACTGCTGCTGCTGCTG 2286  
Qy 2243 TGCAAGTGTCTCGGAGAGATAGCAAGGCGCAAGTCCCAAGTCTTTTAGAGACAGGTCAC 2302  
Db 2287 TGGGTGTCCAGAGGCGCCCATGGGCACTGCGAAGGCTTCAAGAGCCAGCTCTC 2346  
Qy 2303 CACCTCTCTGACCTCCAGCCATACATGGGCGCAGTTCTTAAAGCATCTGCGAGATTCAGA 2362  
Db 2347 TACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGTGGTCTACCTGCGAG-----GA 2400  
Qy 2363 TGGCAGTGTGATGAGGAATCTCGGTTGTCTCAGCAGAGCATCTCTATGATGAGAGCAG 2422  
Db 2401 GACAGCGCGCTGAGGATGCGGTCTCATCAGCAGAGCTCTCTCCCTGGAATGAGGCCAG 2460  
Qy 2423 CAGCAGCTCTTTGACTTCTCTGCACTTCTCGGCTCAGCTGTCTGTAAGATTTGGTGA 2482  
Db 2461 CAGTGGCTCTTTCAGAGTCTTCTCTACGCTTCTGTCAGCAGCAGCGCGTCTCAGGG 2520  
Qy 2483 CAGGTGTATACGATGTCAGGGGATCCCCCAGGGCTCCAGCCTATCCACCGCTGCTG 2542

2521 CAAGTCCTACGTCAGTCCAGGGATCCGAGGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCGTGTTTCGAGACATGGAACAAAGCTGTTCTCAGGTGCGAGGGATGGGTT 2602  
2581 CAGCTGTGTCAGCGCATGAGAACAAAGCTGTTTCGGGGATTCGGCGGACGGCT 2640  
2603 GCTTTTACGTTTGTGTGATGCTTTCTGTGTGACGCTCAGCTTGACCAAGCAAAAC 2662  
2641 GTCCTGCGTTTGTGTGATGCTTTCTGTGTGACCTCAGCTCAGCAAGCAAAAC 2700  
2663 CTTCTCAGCAACCTGCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2722  
2701 CTTCTCAGCAACCTGCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2760  
2723 GACAGTGTGATCTTCCCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2782  
2761 GACAGTGTGATCTTCCCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2820  
2783 GCTGCTCAGCTGCTTCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2842  
2821 GCGGCCCCAGGCTTCTGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2880  
2843 CTTCTGATCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2902  
2881 GACAGCGCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 2940  
2903 TGTCTTCAAGCTGGGAGAACCATCGGAAACAGCTTCTGATGATGGTGCATGATAAATTCAGAA 2962  
2941 CGGCTTCAAGCTGGGAGAACCATCGGAAACAGCTTCTGATGATGGTGCATGATAAATTCAGAA 3000  
2963 TCAGGCTTCTGATCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3022  
3001 TCAGGCTTCTGATCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3060  
3023 CAAGATCTTCTGATCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3082  
3061 CAGATCTTCTGATCTACTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3120  
3083 CAGGCTGTTAGGAGAACCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3142  
3121 TCAGCAAGTTTGGAGAACCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3180  
3143 CTGCTATGCTATCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3193  
3181 CTGCTATGCTATCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3240  
3194 TGGCTCTTCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3253  
3241 CGGCTCTTCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3300  
3254 GGTGCTATGCTATCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3313  
3301 GACTGACACCTGTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3360  
3314 GGTGCTGTCATCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3373  
3361 GGTGCTGTCATCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3407  
3421 ACTGCTCTCAGGTATGTCAGGCTTCTGATGATGGTGCATGATAAATTCAGAA 3454

SULT 7

-10-054-295-224

Sequence 224, Application US/10054295

Publication No. US2003004953A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: NO. US2003004953A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 224:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
/note= "human telomerase reverse transcriptase (hTERT) catalytic protein component"  
SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-10-054-295-224  
Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCCGGCTTGAACAATGACCCCGCGCTCTCTGTCGCCGCGGTGCGTCTCTGTCGCG 82  
Db 40 CCGGCCACCCCGCGATGCGCGCGCTCCCGCTCGGAGCGGTGCTCCCTGCTGCG 99  
QY 83 CAGCGCATACCGGAGGTGTGGCGCTGGCAACTTTGTGCGCGCTGGGCCCGAGG 142  
Db 100 CAGCACTACCGGAGGTGTGGCGCTGGCAACTTTGTGCGCGCTGGGCCCGAGG 159  
QY 143 CAGCGGCTTGTGAACACCCGGGACCCGAGAGATCTACCGCACTTTGTTGCCAATGCCT 202  
Db 160 CTGCGCGCTGTGTCAGCGCGGGACCCGCGGCTTTCCGCGCGCTGTGGCCAGTGCT 219  
QY 203 AGTGTGATGATGCGGCTCAGAGCTCCACCTGCGGACCTTCTCTCCACCAAGTGTC 262  
Db 220 GGTGTGCGTGGCTGGGACGACGCGCGCCCCCGCCCCCTCTCTTCCGCGAGTGTC 279

263 ATCCCTGAAGAGCTGTGGCAGGTTGTGACAGACTCTGGAGCGCAACGAGAGAA 322  
280 CTGCTTGAAGAGCTGTGGCAGGTTGTGACAGACTCTGGAGCGCGGCGGAGAA 339  
323 CGTCTGGCTTTTGGCTTTGAGTGTCTTAACAGAGCCAGAGCGGCGCTTCCCATGGCCTT 382  
340 CGTCTGGCTTTGGCTTTGAGTGTCTTAACAGAGCGGCGGCGGCGGCGGCGCTT 399  
383 CACTAGTAGCTGTGAGTGTCTTAACAGAGCGGCGGCGGCGGCGGCGGCGGCGCTT 442  
400 CACACAGAGCTGTGAGTGTCTTAACAGAGCGGCGGCGGCGGCGGCGGCGGCGCTT 459  
443 TGCAATGATGCTACTGTGTAGCGAGTGGGCGGAGCGGCGGCGGCGGCGGCGGCGCTT 502  
460 GGCCTGGGCGGCTGTGCTGGCGCGTGGGCGGAGCGGCGGCGGCGGCGGCGGCGCTT 519  
503 CTGTGCTCTTTTATCTTGTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCTT 562  
520 CTGGCGCTCTTTTGTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCTT 579  
563 GTACCAATTTGTGCCACACAGGATATCTGGCGGCTGTGCTGGCGGCGGCGGCGGCGCTT 622  
580 GTACAGCTGGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCTT 639  
623 CGACCCGCTGGGCGGAGTTTCACTAACCTTAGGTTCTTAAACAGATCAAGAGCAGTAG 682  
640 GCCTCTGGGATGC-----GAAACGGGCGTGGAAACATAG 672  
683 TCGCAGAGAGCCCGGAAACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 742  
673 CGTCAAGGAGCGGCGGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 732  
743 TCTCAGCAGTACAGTGTGCTTTCAGTAAAGAGCGGCGGCGGCGGCGGCGGCGGCGCTT 802  
733 CAGTGGCCAGCGAGTCTGCGGTTGCCAGAGGCGGCGGCGGCGGCGGCGGCGGCGCTT 792  
803 GGAGGAGGAGCCCGCAGGAGTGTCTCAACCCGATCAGGCAATATGAGGTGCCAAG 862  
793 GGAGCGAGCCCGGTTGGGCGGAGGCTTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 852  
863 TCCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 913  
853 TGACCGTGGTTTCTGT 912  
914 AGAAGAGTGTGTGAGT 970  
913 GGGTGGCTCTCTGGCAGCGGCGGCTTCCACCGCATCCGTTGGCGGCGGCGGCGGCGGCTT 972  
971 CTCACAT---CTCTGTGTACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1027  
973 CCCCCATCCAGATCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1032  
1028 TGAGACAGACATTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1087  
1033 CGAGACCAAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1089  
1088 CTTACTCAGCAACTCCAGCTTAACTTGAATGGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 1149  
1148 TCTGGGCTAAGCGCTAGGACATTCAGGAGCACTCTGCGAGGACACACCGCTTATCGCGTCG 1207  
1150 TCTGGGTTCCAGGCGCTGGATGCCAGGACTCCCGCGAGTTGCCCGGCTTGGCGGCGGCTT 1209  
1208 ATACTGCGAGATCGGCGGCGGCTTCTTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1267  
1210 CTTACTGCAATCGGCGGCGGCTTCTTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269  
1268 TGTCAAGACTCTCAGGTTCATTTGAGGTTTTCGAAACAGCAACCAAGGTCAGAGTGC 1327  
1270 CGGGTGTCTCTCAAGAGCACTGCGCGGCTGAGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1329

1328 CT-----TGAACACAGGCC 1342  
1330 CTGTGCCGGGAGAACCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGACAGAGCC 1389  
1343 ACCGACCTCATGAGATTTGCTCCGCTGTGACAGAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CGTCTGGCTGTGTGAGTGTCTCCGCGCAGCAGCAGCGGCTTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGGCTGTCTCTGGAAGT 1462  
1450 GCGGGCTGTCTCTGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1509  
1463 CGCTCTCTTAAAGACTTAAAGAGTTCATCTCGTTGGGAAATACGCGAGCTATCACT 1522  
1510 CGCTCTCTCAGAGACACAGAGAGTTCATCTCTCTGGGAGCATGCCAGCTCTGCT 1569  
1523 GAGAGACTGTGTGGAAGATGAAAGTAGAGATTCGCTCTGGCTCCGCGAGAGCCGGG 1582  
1570 GCAGGAGCTGAGCTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGAGCCAGG 1629  
1583 GAAGACCGTGTCCCGCTGCGAGACCGCTCTGAGGAGAGGATCTGGCTAGGTTCT 1642  
1630 GGTGGCTGTGTCTCGCGCGCAGAGACCGCTCTGCGTAGAGATCTCTGGCAAGTTCT 1689  
1643 GTTCTGGCTGATGAGACACATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1702  
1690 GCATGGCTGT 1749  
1703 GAGCATTCTCAGAGAGAGAGCTCTTCTTCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1762  
1750 GACCACTTCAAAAGACAGGCTCTTCTTCTAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1809  
1763 GAGCATTGTGAGTCAAGCAACACTTGTAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1822  
1810 AGCATTGTGATCAGACAGCACTTGAAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1869  
1823 GGTGAGCATCAGAGAGACCTGT 1882  
1870 GGTGAGCAGCATCAGAGAGAGCGGCGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1929  
1883 CAAGCCCAAGCGCTGTGGCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1942  
1930 CAGGCTGTGAGGCTGTGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1989  
1943 GGGCAG 2002  
1990 CGCAG 2049  
2003 CAACTGTGAGCGGACAAACATCTCTCACTTATGGGCTTCTGTGTGTGTGTGTGTGTGTGT 2062  
2050 CAACTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTT 2109  
2063 CATCTACAG 2122  
2110 TATCAG 2169  
2123 GATGTACTTGT 2182  
2170 GCTGTACTTGT 2229  
2183 GGT 2242  
2230 CAGGAGAGT 2286  
2243 TGCAGT 2302  
2287 TGGCGT 2346  
2303 CACCTCTCTGT 2362  
2347 TACCTGT 2400  
2363 TGCCAGT 2422

```
|||||
2401 GACAGCCCGCTGAGGATGCGTGTCTATCGAGCAGAGCTCTCCCTCAATGAGGCCAG 2460
2423 CAGCAGCCTGTTGACCTTCTCTGCACTTCTCGCTGTCAGTGTGTAAGATTGGTGA 2482
2461 CAGTGGCCTCTTCGACGCTCTCTAGCTTCTATGTGCCACCGCGTGCAGCATCAGGG 2520
2483 CAGTGTCTATACGAGTCAGAGGATCCCGAGGCTCCAGCTATCCACCTCTCTG 2542
2521 CAAGTCTCTGTCAGTGCAGGGATCCCGAGGCTCCATCTCTCCAGCTCTCTG 2580
2543 CAGTCTGTTTCCGAGACATGAGAACCAAGCTTGTGTGAGGTGCAGCGGATGGGT 2602
2581 CAGCCTGTGTACGGCAGACATGAGAACCAAGCTTGTGCGGATTTGCGCGGACGGCT 2640
2603 GCTTTTACGTTTGTGATGACATTTCTGTGTGAGCTCTCACTTGGACCAAGCAAAAC 2662
2641 GCTCCTGCTTGTGTGATGATTTCTTGTGTGACCTTCACTTCACTTCACTTCACT 2700
2663 CTTCCTCAGCACCTTGTCTCATGCGCTTCTGAGTATGGTGCATGATAAACTTCAGAA 2722
2701 CTTCCTCAGCACCTTGTCTCGAGGTGTCCTGAGTATGCTGCTGGTCACTTCGGAA 2760
2723 GACAGTGTGAACTTCCCTGTGAGCCTGTACCTTGGTGTGAGCTTCCATACAGCT 2782
2761 GACAGTGTGAACTTCCCTGTGAGCCTGTACCTTGGTGTGAGCTTTCAGT 2820
2783 GCTGTCTCACTGCTGTTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2842
2821 GCGGCCCAAGGCTATTTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2880
2843 GTTCTGTGACTCTCAAGTTATGCCAGACCTCAATTAAGACAGCCTCACTTCCAG 2902
2881 GCAGAGGACTACTCCAGCTATGCCGAGCTTCCATCAGAGCAGCTTCACTTCAAC 2940
2903 TGTCTTCAAGCTGGGAGACCATGCGGACAGCTTCTGCTGCTGCTGCTGCTGCTG 2962
2941 CGGCTTCAAGCTGGGAGAACATGCTGCAACTCTTTGGGGCTTGGGCTGAGTG 3000
2963 TCAGGCTTATTTCTAGACTTGAGGTGAGCAAGCTTCCAGACAGCTGCTCAATATA 3022
3001 TCAGAGCTTCTGATTTGAGTGTGAGGTGAGCAAGCTTCCAGAGCTGCTCAATATA 3060
3023 CAAGATCTTCTGCTTCAAGCTTCAAGTGTGATGATGATGATGATGATGATGATGAT 3082
3061 CAAGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
3083 CCAGCTGTTAGGAAGAACTCACATTTCTTGGGCATCATCTCCAGCAAGCATCTG 3142
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTCTGACAGGCTTCCCT 3180
3143 CTGCTATGCTATCTGAGGTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 3193
3181 CTGCTATCTTCTGAGGTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 3240
3194 TGGCTCTCTTCTCTGAGCCGACATTTGCTGCTTACAGGCTTCTGCTTCAAGCT 3253
3241 CGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
3254 GGCTGCTCATCTGCTATCAAAATGCTTCTTGGGACCTTCTGAGGACAGCCAAAAC 3313
3301 GACTCGACACCGTGTCACTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
3314 GCTGTGCGGAGCTTCCAGAGCGCAATGACCATCTTAAAGTGTGAGTGTGAGTGTG 3373
3361 GCTGAGTGTGAGCTTCCGAGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
3374 CCTAAGCAGACTTTTCAAGACATTTTGAATAA 3407
3421 ACTGCCCTCAGACTTCAAGACATCTCTGACTGA 3454
```

```
US-10-054-611-224
; Sequence 224, Application US/10054611
; Publication No. US20030059787A1
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
;               Lingner, Joachim
;               Nakamura, Toru
;               Chapman, Karen B.
;               Morin, Gregg B.
;               Harley, Calvin
;               Andrews, William H.
; TITLE OF INVENTION: No. US20030059787A1el Telomerase
; NUMBER OF SEQUENCES: 225
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Townsend and Townsend and Crew LLP
;   STREET: Two Embarcadero Center, 8th Floor
;   CITY: San Francisco
;   STATE: California
;   COUNTRY: United States of America
;   ZIP: 94111
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Patent In Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/054,611
;     FILING DATE: 18-Jan-2002
;     CLASSIFICATION: 536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 08/854,050
;       FILING DATE: <Unknown>
;       APPLICATION NUMBER: US 08/846,017
;       FILING DATE: 25-APR-1997
;       APPLICATION NUMBER: US 08/844,419
;       FILING DATE: 18-APR-1997
;       APPLICATION NUMBER: US 08/724,643
;       FILING DATE: 01-OCT-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Apple, Randolph T.
;       REGISTRATION NUMBER: 36,429
;       REFERENCE/DOCKET NUMBER: 015389-002930US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (415) 576-0200
;       TELEFAX: (415) 576-0300
;     INFORMATION FOR SEQ ID NO: 224:
;       SEQUENCE CHARACTERISTICS:
;         LENGTH: 4015 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;         MOLECULE TYPE: cDNA
;         FEATURE:
;           NAME/KEY: CDS
;           LOCATION: 56..3454
;           OTHER INFORMATION: /product= "hTERT"
;           /note= "human telomerase reverse
;             transcriptase (hTERT) catalytic protein
;             component"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 224:
US-10-054-611-224
```

```
Query Match      45.8%; Score 1600.8; DB 15; Length 4015;
Best Local Similarity 68.9%; Pred. No. 0;
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCCGCCCTTGAGCAATGACCGCGCTCTCTGCTTGGCCCGCGGTGGCTCTCTGCTGCG 82
Db 40 CCCGCCACCCCGCGATGCCGCGCTTCCCGCTGCGAGCCGCTCCCTGCTGCG 99
QY 83 CAGCGATACCGGAGGTGTGCGCTGGCAACCTTTGTGCGGCGCTGGGGCCCGAGGG 142
Db 100 CAGCCACTACCGGAGGTGTGCGCTGGCCACGTTCTGCGGCGCTGGGGCCCGAGGG 159
```

143 CAGCGCGCTTGTCAAACCCGGGACCCGAAGATCTACCGCACTTGTGTTGCCCAATGCT 202  
160 CTGGCGGCTGTGTGCAGCGCGGGACCCGGCGGCTTTCGGCGCGTGTGTGCCAGTGCCT 219  
203 AGTGTGCATGCACCTGGGCTCACAGCTCTCACTCTGCCGACTTTCCTTCACACAGGTGC 262  
220 GGTGTGCGTGCCTTGGGACGACAGCGCCGCCGCCGCCCTCTCTTCGCCAGAGTGC 279  
263 ATCCCTGAAGAGCTGTGGCCGAGGTTGTGCAGAGACTCTGCAGAGCGCAACGAGAGAA 322  
280 CTGCCTGAAGAGCTGTGTGGCCCGAGTGTCTCAGAGGCTGTGCAGCGCGCGCGCGAAGAA 339  
323 CGTGTGCGCTTTTGGCTTTGAGCTGCTTAACAGAGCCAGAGCGGGGCTCCCATGGCCCT 382  
340 CGTGTGCGCTTTCGGCTTCGCGCTGTCTGGAAGGGGCGCGCGGGGCCCCCGAGAGCCCT 399  
383 CACTAGTAGCGGTGAGTACTTGTGCCCAACACTGTTATTGAGACCCCTGCGTGTCAATGG 442  
400 CACCACCAAGCGGTGCGCACTACTTGCCTCAACACAGGTGACAGCACTGTGGGGGAGCGG 459  
443 TGCATGGATGCTACTGTTGAGCGAGTGGGCGAGCACTGCTGTCTTACCTGTGTGCACACA 502  
460 GCGTGTGGGCTGCTGCTGCGCGCGTGGGCGAGCACTGTGTGTTTCACTGTCTGGCAGC 519  
503 CTGTGCTCTTTATCTTCTGTGTGCCCCCGCAGCTGTGCCCTACACAGGTGTGTGGGTCTCC 562  
520 CTGCGCGCTCTTTGTGTGTGTGCTCCCACTGCGCTTACCAAGTGTGTGGGCGCGCGCT 579  
563 GTACCAAAATTTGTSCACACAGGATATCTGGCCCTCTGTGTCGCTAGTTTACAGGCCAC 622  
580 GTACCAAGCTCGCGGTGTGCACTCAGCGCCCGCCCCCGCACACACTAGTGGACCCCGAAG 639  
623 CGACCCGCTGGGCAGGAATTCCTAACTTAGTTCTTTCAACAGATCAGAGCAGTAG 682  
640 GCGTCTGGAGTGC-----GACGGGCTTGGAAACCATAG 672  
683 TCGCCAGGAGACCGAAACCGCTGGCTTGCATCTCGAGTACAAGAGGCATCTGAG 742  
673 CGTCAGGAGGCGCGGGTCCCGCTGGGCTCCGACGCCGGGTGCGAGAGCGCGCGGG 732  
743 TCTCACCAAGTCAAGTGTGCTTCACTTAAAGAGCCAGATGCTATCCTGTCCCGAGAT 802  
733 CAGTGCAGCGCAAGTCTGCGTTCGCCCAAGGCCCCAGGCGTGGCGCTGCCCTGAGCC 792  
803 GGAGAGGGAACCCACAGCAGGTGCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 GGAGCGGACGCCGTGTGGCAGGGTCTCTGGGCCACCCGGGCGAGGACGCTGTGACCGAG 852  
863 TCTGTGCTCGGTCCCGCAGGTGCGCTACT-----GCAGAGAAAGATTGTCTTCTAA 913  
853 TGACCGTGTGTTCTGTGTGTGTCACTTCCAGACCCGCCGAAGAGCCACCTCTTTGGA 912  
914 AGGAAAGTGTCTGACTGAGTCTCTC---TGGGTGGTGTGCTGTGTAAACAAAGGCCAG 970  
913 GGGTGGCTCTCTGACACGGCACTTCCACCCATCCGTGGGCGCGCGAGCACCAAGCGGG 972  
971 CTCCACAT---CTCTGCTGTACACCCCGCCAAATGCTCTTTCAGTCTCAGGCCATTTAT 1027  
973 CCCCCCATCCACATCGCGGCCACCAAGTCCCTGGGACAGGCTTGTCCCCGGTGTACGC 1032  
1028 TGAGACCGAGCATTTCTTTTACTCAGGGGAGATGGCCAGAGCGTCTTAAACCCCTCAT 1087  
1033 CGAGACCGAGCACTTCTCTACTCTCTCAGGGGA---CAAGAGAGAGTGTGGGCCCTCCT 1089  
1088 CTTACTCAGCAACCTCCAGCTTAATCTCACTGGGCGCAGGAGACTGTGTGGAGATCATCT 1147  
1090 CTTACTCAGCTCTCTGAGGCCCGAGCTGACTGTGGCTCGAGGCTGTGTGAGACCATCT 1149  
1148 TCTGGCTCAAGGCTTAGGACATCAGGACCACTCTGCAAGCACACCGTCTATCCGCTCG 1207  
1150 TCTGGGTTCCAGGCCCTGATGCCAGGACTCCCCGCAAGTTGCCCCGCTGTGCCCGCAGC 1209

Qy	1208	ATATCGGCAGATGGGCCCCCTGTTCCAA	CAGCTGCTGTGTGAACATGCAAGTCCCAATA	1267
Db	1210	CTACTGGCAAAATGGGCCCCCTGTTTCT	TGGAGTCTGTTGGGAACACGCGAGTCCCCCTA	1269
Qy	1268	TGTCAGACTCTCTCAGGTCACTTGCAG	GTTCGGAACAGCAAAACAAAGGTGACAGATGC	1327
Db	1270	CGGGTGTCTCTCAAGACGACATGCCG	CTGCGCTGCGGTCACCGACGACGCGTGT	1329
Qy	1328	CT-----	-----TGAACACACGACCC	1342
Db	1330	CTGTGCCGGGAGAACCCACAGGCTCT	GTGTGGCGGCCCCCGAGGAGGAGGACACAGACCC	1389
Qy	1343	ACCGCACCTCATGATTTGCTCGCCTGC	ACAGCACTCCCTGGCAGGTATATGTTTCT	1402
Db	1390	CGTCGCTGTGTGAGCTGTCTCGCAG	ACACAGCAGCCCCCTGGCAGGTGACGGTTCGT	1449
Qy	1403	TGCGCCCTGTCTCTGCAAGTGTGTCT	GTCTGTCTGTGGGTACCAAGCAATATGAGCG	1462
Db	1450	GGCGCCCTGCTCGCGCGCTGTGTGCC	CCAGCGCTCTGGGCTCAAGCAACAGAACG	1509
Qy	1463	CGCTTCTTTAAGAACTTAAGAAAGTT	CACTCTGTTGGGAAATACGCAAGCTATCACT	1522
Db	1510	CGCTTCTCTCAGGAACACCAAGAAAG	TTCATCTCCTTGGGAAGCATGCCAAGCTCTCGCT	1569
Qy	1523	GCAGAACTGATGTGGAAGATGAAGGT	TAGAGATTGCCACTGGCTCCGACAGCCCGGG	1582
Db	1570	GCAGAGCTGACGTGGAAGATGAGCGT	GGGACCTGGCTGGCAGAGCCCCAGG	1629
Qy	1583	GAAGACCGTGTCCCGCTGCAGACAC	CGTCTGAGGAGAGGATCTGGCTACGTTCT	1642
Db	1630	GTTGGCTGTGTCCGCGCAGACACCGT	CTGTGCTGAGGAGATCTCTGGCCAAAGTTCT	1689
Qy	1643	GTTCTGGCTGATGACACATACCTGGT	TACAGTGTCTAGTCACTTTTATCATCAACA	1702
Db	1690	GCACCTGGCTGATGAGTGTGTAAGT	GTCTGAGCTGTCTCAGTCTTCTTTATGTCA	1749
Qy	1703	GAGCACATTCAGAAAGAACAGGCTCT	TCTTCTTACCGTAAGAGTGTGTGGCAAGCTGCA	1762
Db	1750	GACCACGTTTCAAAGAAACAGGCTCT	TTTTTCTACCGAAGAGTGTCTGGACAAAGTTGCA	1809
Qy	1763	GAGCAATCGAGTCAGGCAACACCTTC	GTGAGAGTGGGCTACGGGAGCTGCACAAGAGGA	1822
Db	1810	AAGCATTTGAAATCAGACAGACATTT	GAAAGAGGTGCAGCTGGGAGCTGTCCGAAGCAGA	1869
Qy	1823	GGTCAGGCATCACAGGACACCTGGCT	TAGCCATCTGCCATCTGCAGACTCGCTTCATCCC	1882
Db	1870	GTCAGGCAGCATCGGGAACGACGGC	CCGCCCTGTCTGACGTCAGACTCGCTTCATCCC	1929
Qy	1883	CAAGCCCAACGGCTCGGCCCATTTGT	GAACATGAGTTATAGCATGGGTACGAGAGCTTT	1942
Db	1930	CAAGCCTTCACGGGCTGCGGCCGAT	TTGTGAACATGGACTACGTGTGGGAGCCAGAAAGTT	1989
Qy	1943	GGGCAGAAAGGAGCAGAGCCGACGAT	TTTCAACCAGCGTCTCAAGACTCTCTTTCAGCATGCT	2002
Db	1990	CCGACAGAAAGAGGGCCGAGCGCT	CTCACTTCGAGGTTGAAGGCACCTGTTTCAGCGTGT	2049
Qy	2003	CAACTATGAGCGGCAAAACAATCCT	CACCTTATGGGTCTTCTGTACTGGGTATGATGA	2062
Db	2050	CAACTACAGAGCGGCGCGGCCCTCT	CTGTGGCGCTCTGTGCTGGGCGCTGGACGA	2109
Qy	2063	CATCTACAGGACCTGGCGGCGCTTT	GTGTGGCGTGTGGCTGTCTGGACAGACACCCAG	2122
Db	2110	TATCCACAGGGCTTGGGACCTTCGT	GTGTGGCTGTGGGCGCCACAGACCCGCGCCTGA	2169
Qy	2123	GATGTACTTTGTTAAGCAGATGTGAC	CGGGGCGCTATGATGCGCATCCCCCAGGGTAAAGCT	2182
Db	2170	GCTGTACTTTGTCAAGGTGGATGTGA	CGGGCGGTACGACACCATCCCCCAGGACAGGCT	2229
Qy	2183	GGTGGAGGTTGTTGCCATATGATCAG	GCACCTCGGAGACAGTACTGTATCCGCGAGTA	2242
Db	2230	CACGAGGTCATGCCAGCATCATCA	---AACCCGAGAAACAGTACTGCTGCTGCGTA	2286
Qy	2243	TGCAGTGTGTCGGAGAGATAGCAAG	GCCAAGTCCCAAGTCTTTTAGGAGACAGGTCAC	2302

2287 TGCCGTGTCAGAGGCGCCGATGGGACGTCGCGAGGCGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGACCTTCAGCCATACATGGGCGAGTCTCTTAAGCATCTGAGAGATTGAGA 2362  
2347 TACCTTGACAGACCTCCAGCGGTACATGCGACATGTCGTGGCTCACCTGCAG-----GA 2400  
2363 TGCCAGTGCATGAGGAATCCGTTGTCATCGAGAGAGCATCTATGAATGAGAGAG 2422  
2401 GACCAGCCCGTGAGGATGCGCTGTCATCGAGAGAGCTCTCCTGAAATGAGGCCAG 2460  
2423 CAGAGCGCTGTTGACATCTTCCTGCACTTCCTGCGTCAAGTGTGCTAAAGATTGGTGA 2482  
2461 CAGTGGCTCTCTGACGCTTCTTACGCTTCTATGTGCCACCGCGTGGCATCAGGGG 2520  
2483 CAGTGTCTATACGAGTGCAGGGCATCCCGAGGCTCCAGGCTCCAGCTATCCACCTGCTG 2542  
2521 CAGTCTCTAGCTGCGAGTGCAGGGGATCCCGAGGCTCCATCTCTCCACGCTGCTG 2580  
2543 CAGTGTGTTTGGAGACATGAGAACAGCTGTTTGTGAGGTGCAGCGGATGGGT 2602  
2581 CAGCTGTGCTACGCGACATGAGAACAGCTGTTTGGGGGATTCGGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACATCTCTGTTGGTACGCTCTCCTGAGCAAGCAAAAC 2662  
2641 GCTCTGCGTTTGTGATGATTTCTGTTGGTGACACCTCACTCACCACCGCGAAAC 2700  
2663 CTTCCTCAGCACCTGCTCCATGGCTTCTGATGATGGGTGCATGATAAACTTCAGAA 2722  
2701 CTTCCTCAGGACCTGCTCGAGGTGCTCCCTGATGATGCTGGTGGTAACTTCGGAA 2760  
2723 GACAGTGGTGAATCTCCTGTGAGCCTGTGACCTGGGTGTGAGCTCCATACAGCT 2782  
2761 GACAGTGGTGAATCTCCTGTGAGACGAGGCGCTGGGTGCGACGGCTTTTGTTCAGAT 2820  
2783 GCCTGCTCACTGCTGCTTCTCCTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 2842  
2821 GCGGCGCCAGCGCTATTCCTCTGTGTGGCGCTGTGCTGTGATATCCGAGCCTCGAGGT 2880  
2843 GTTCTGTGACTCTCAGGTTATGCCAGACCTCAATTAAGACGAGCCTCACTCCAGAG 2902  
2881 GCAGAGCGACTACTCCAGCTATATCCCGGACCTCCATCAGAGCGAGTCTCACCTCAAAC 2940  
2903 TGTCTTCAAAGCTGGGAGACCATGCGGACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2962  
2941 CGGCTTCAAGGCTGGGAGGACATGCTGCGAACTCTTTGGGGTCTTGGGCTGAGGTG 3000  
2963 TCACGCTTATTTCTAGACTTGCAGTGAACAGCCTCCAGACAGTCTGCATCAATATATA 3022  
3001 TCACAGCCTGTTTCTGATTTGTCAGGTGAACAGCCTCCAGACGCTGTCACCAATCTA 3060  
3023 CAAGATCTTCTGCTTCAAGCCTCAGAGTTCATGATGATGATGATGATGATGATGATGATG 3082  
3061 CAAGATCTTCTGCTGAGGCTGACAGGTTTCAACGATGATGATGATGATGATGATGATGAT 3120  
3083 CCAGCGTGTAGGAAGAACTCACATCTTCTTGGGCAATCATCTCCAGCAAGATCTG 3142  
3121 TCAGCAAGTTTGAAGAACCCACATTTTCTGCGCTCATCTCTGACACGCGCTCCCT 3180  
3143 CTGCTATGCTATCTCAGAGTCAAGATCCAGGAATGACATTAAG-----GCTC 3193  
3181 CTGCTACTCAATCTGAAAGCCAAAGAACAGAGGATGTCGCTGGGGGCAAGGGCGCGC 3240  
3194 TGCTCTCTTCTCTCTGAGCGCGCATTTGGCTCTGCTACAGGCTTCTGCTCAAGCT 3253  
3241 CGGCTCTCTGCTCTCGAGGCTGTCAGTGGCTGTGCGCACCAAGATTTCTGCTCAAGCT 3300  
3254 GGCTGCTATCTGCTCATCTAATAATGTTCTCTGGACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTCGACACCGGTGTCACCTAGTGCCTCTCTGGGTCACTCAGGACAGCCCAAGCGCA 3360  
3314 GCTGTCCGGAAGCTCCAGAGCGGACATGACATCTTAAAGCTGACGATGCCAGC 3373

Db 3361 GCTGAGTGGAGAGCTCCCGGGAGCAGCGCTGACTGCTCCCTGGAGGCGCGAGCCACCCGGC 3420  
Qy 3374 CTTAAGCAGAGCTTTTCAGACCATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGAGCTTCAAGACCATCTCTGACTGA 3454  
RESULT 9  
US-10-105-963-1  
; Sequence 1, Application US/10105963  
; Publication No. US20030068818A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Denning, Chris  
; APPLICANT: Clark, A. John  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
; TITLE OF INVENTION: Recombination  
; FILE REFERENCE: 731/002  
; CURRENT APPLICATION NUMBER: US/10/105,963  
; PRIOR FILING DATE: 2002-03-21  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION:  
US-10-105-963-1

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

Qy 23 CCGCGCTTTGACACAAATGACCCGCGCTCTCTGTTGCCCGCGGTGGCTCTCTGCTGCG 82  
Db 40 CCGCGCACCCCGCGATGCGCGGCTCTCCCGCTGCGAGCCGTGCTCTCTGCTGCG 99  
Qy 83 CAGCCGATACCGGAGGTGTGCGCTGGCAACCTTTGTGCGGCGCTTGGGCGCCGAGGG 142  
Db 100 CAGCCACTACCGAGAGTGTGCGCTGGCCACCTTGTGCGGCGCTGCGGCGCCGAGGG 159  
Qy 143 CAGCGGCTTGTGCAACCCGCGGAGCCGAGATCTACGACATTTGCTTGGCCCAATGCT 202  
Db 160 CTGCGGCTTGTGCGAGCGCGGAGCCCGCGGCTTCCGCGCGCTGCTGCGCGCTGCT 219  
Qy 203 AGTGTGATGACATCGGCTCACAGCTTCCACCTTCCGAGCTTTCCTTCCACCGAGTGT 262  
Db 220 GGTGTGCTGCTGCGTGGAGCGCACGCGCGCCCGCGCTTCTTCCGCGAGTGT 279  
Qy 263 ATCCCTGAAGAGCTGTGCGGAGGCTGTGCGAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 280 CTGCTTGAAGAGCTGTGCGGAGGCTGTGCGAGGCTGTGCGAGCGCGCGCGCGAGAA 339  
Qy 323 CGTGCTGCTTTTGGCTTTGAGCTTAAACGAGCGCAGAGGCGGCTTCCCATGCGCTT 382  
Db 340 CGTGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 399  
Qy 383 CACTAGTAGCGTGGTACTTATGCTTATGAGCCCAACATGTTTATGAGCCCTGCTGCTG 442  
Db 400 CACCAACAGCGTGGCGAGTACTTACCTGCGCAACACCGTACCGACGCACTGCGGAGCG 459  
Qy 443 TGCTAGGATGCTACTTGTGAGCGGAGTGGCGAGCTGCTGCTTACCTGCTGCTGCTG 502  
Db 460 GCGTGGGGGCTGCTGCTGCGCGCGGCGGAGCGTGTGCTGCTGCTGCTGCTGCTGCTG 519  
Qy 503 CTGTGCTCTTATCTTCTGCTGCGCGCGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTG 562

520 CTGCGCGCTTTGTGTGGTGGTCCGACCTACGAGTGTGGGCGCCCGCT 579  
563 GTACCAATTTGTGCCACCAAGATATCTGGCCCTCTGTCTCGCTAGTTACAGGCCAC 622  
580 GTACAGCTCGGCGCTGCCACTCAGGCGCGGCCCGCCACACGCTAGTGGACCCGAG 639  
623 CGAGCCGTGGGCGAGGAATTTCACTAACCTTAGTTCTTACAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGGATGC-----GAAACGGCGCTTGAACCATAG 672  
683 TCGCCAGGAAGCACCGAAACCCCTGGCTTGCATCTCGAGGTACAAAGAGGATCTGAG 742  
673 CGTCAGGAGGCGCGGCTCCCTCTGGCTTGCACGACCCCGGCTGCGAGAGAGCGCGGG 732  
743 TCTCACAGTACAGTGTGCTTCAAGTACAGAGGCGAGATGATCTCTGTCTCCAGAGT 802  
733 CAGTGCACAGCGAAAGTCTGCGCTTTCGCAAGAGGCGCGGCTGCGGCTGCGCTGAGCC 792  
803 GGAGAGGAGCCACACAGGCGAGGTGCTACCAACCCCATCAGGCAATCATGCGTGCCAG 862  
793 GGAGCGGAGCGCGTGTGGGAGGAGTCTGGGCCACCCCGGCGAGCGGTGGACCGAG 852  
863 TCCTGCTCGGTCCCGGAGTGCCTACT-----GCAGAGAAAGATTTGTCTTTCTAA 913  
853 TGACCGTGGTTCTGTGTGTGTGTCACTGCGACACCGCGCAAGAGGCCACTCTTTGA 912  
914 AGGAAGGTGTGACCTGTGCTCTC-----TGCGTGGTGTGTGTGTAAACACAGGCCAG 970  
913 GGTGCGCTCTCTGGACGCGCCACTCCACCCATCCGTTGGGCGCGCAGCACCGCGG 972  
971 CTCACAT---CTGCTGTGTACACACCGCGCAAAATGCTTTTTCAGCTCAGGCGCAATTAT 1027  
973 CCCCCATCCACATCGCGGCCACACACCTGCTGGGACACGCTTGTCCCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGAGAGATGGCCAGAGAGGTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGACATTTCTTACTCTCTCAGGCGA---CAAGGAGCAGCTGCGGCCCTCT 1089  
1088 CTAATCTCAGCAACTCCAGACCTTAATCTTGAATGGGCGCAGAGACTGGTGGAGATCATTT 1147  
1090 CTAATCTCAGCTCTGTAGGCGCCAGCTGACTGTGGCTCGAGGCTGTGGAGACCATTT 1149  
1148 TGTGGCTCAAGGCTTAGGACATCAGACACCTCTGAGAGACACACGCTATCGGCTG 1207  
1150 TGTGGGTTCCAGGCGCTGGATSCCAGGAGCTCCCGCAGGTTGCGCGCTGCGCCAGCG 1209  
1208 ATACTGGCAGATGGGCGCTGTTCACACAGCTGCTGTGTAACCATGACAGGTGCCAATA 1267  
1210 CTACTGGCAATGCGGCCCTGTCTTGGAGCTGCTTGGGAACACGCGCAGTGGCCCTTA 1269  
1268 TGTGAGATCTCTCAGGTCAATGCAAGTTCGAAACAGCAACCAACAGGTGACAGATGC 1327  
1270 CGGGGTGCTCTCAAGAGCGACTCCCGCTCGAGTGCCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACAGAGCC 1342  
1330 CTGTCCCGGAGAGCCCGAGGCTGTGGGCGGCCCGGAGGAGGAGGACAGACCC 1389  
1343 ACCGACCTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
1390 CGGTGCTGTGAGTGTCTCCGACAGCAGCAGCGCCCTGGCAGGTGTACGGCTTGT 1449  
1403 TGGGCTGTCTGCAAGTGTGTGTCTGTGTCTGGGTACAGGCAACATGAGCG 1462  
1450 GGGGCGCTGCTGCGCGGCTGGTGGTGGCCCGCAGGCTCTGGGCTCCAGGCAACAGCG 1509  
1463 CGCTTTCTTAAGAACTTAAAGAGTTCATCTGTTGGGAAATACGGCAAGTATCACT 1522  
1510 CGCTTCTCAGAACACCAAGAGTTCATCTCCCTGGGAGGATGCGAAGTCTCGT 1569  
1523 GCAGAACTGATGGAGATGAAGTAGAGATGTCATGCTGCGCGAGAGCCCGG 1582

1570 GCAGGAGCTGACCTGGAAGATGAGCGTGGCGGAGCTGCGCTTGGCTGCCAGAGGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGCAGAGCACCGTGTGAGGAGAGGATCTTGGGTACGTTCT 1642  
1630 GGTGTGCTGTGTCCCGCGCAGAGCACCGTCTGCTGAGGAGATCTTGGCCAAAGTCTCT 1689  
1643 GTTCTGGCTGATGGACACATAGTGTGACAGTGTAGTGTAGTCACTTTTACATCAGAGA 1702  
1690 GCACTGGCTGATGAGTGTACGTGCTGAGCTGTCTCAGGCTCTTTCTTTTATGTACCGGA 1749  
1703 GAGCATTCCAGAAAGACAGGCTCTTTCTTACCGTAAAGTGTGTGGAGCAAGCTGCA 1762  
1750 GACACGTTTCAAAAGAAACAGGCTCTTTTCTTACCGAAGAGTGTCTGAGCAAGTTGCA 1809  
1763 GAGCATTTGAGTGTGAGCAACACCTTGAAGAGTGGGCTACCGGAGCTGTACAGAGAGA 1822  
1810 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGAGTGGGGAGCTGTGGGAAGCAGA 1869  
1823 GGTGAGCATCACAGGACACCTGCTAGCTAGCCATCTGCAGACTGCGCTTCATCCC 1882  
1870 GGTGAGGAGCTCGGGAAGCCAGGCCCGCTGTGACGTCAGACTCCGCTTCATCCC 1929  
1883 CAAGCCCAACGCTCGGCGCTTGTGAAACATGAGTTATAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCCTGACGGCTCGGCGCTTGTGAAACATGAGTACGTCGTTGGGAGCCAGAACGTT 1989  
1943 GGCAGAGGAAGCAGCGCCAGCATTTTACCCAGCGCTCTCAAGACTCTTTCAGCATGCT 2002  
1990 CCGAGGAAAGAGGCGCGAGGCTCTACCTCGAGGTGAAGGACATGTTTCAGGCTGCT 2049  
2003 CAATATGAGCGGCAAAAACATCTCTCACTTATGAGGCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGCGCGCGCGCTCTCGGCGCTCTGTGCTGGGCTGGGCTGGAGCA 2109  
2063 CATCTACAGACCTGCGGCGCTTGTGCTGCTGTGCTGCTGCGGCTGGACGACACCCAG 2122  
2110 TATCCAGGCGCTGCGCGACCTTGTGCTGCTGTGCTGCGGCGCCAGGACCCCGCTGA 2169  
2123 GATGTACTTTGTTAAGCGAGATGTGACCGGCGCTTATGATGCCATCCCCCAGGGTAAGCT 2182  
2170 GCTGTACTTTGTCAAGTGGATGTGACGGCGCTGACACCATCCCCCAGGACAGGCT 2229  
2183 GGTGAGGTGTGTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2242  
2230 CAGCGAGGTCACTCGCAGCATCATCA---AACCCAGAACACGTAAGTCTGCTGCTGCTG 2286  
2243 TGCACTGCTCCGAGAGATAGCAAGGCCAAGTCCACAAAGTCTTTTAGGAGACAGGTCAAC 2302  
2287 TGCGTGGTCCAGAGCGCGCCATGGGCACTCGGCAAGGCCITCAAGAGCCAGTCTC 2346  
2303 CACCTCTGTGACTCGAGCATACTAGGCGCAGTCTTAAAGCATCTGAGGATTCAGA 2362  
2347 TACTTGACAGACTCCAGCGCTACATGACAGTTCGTTGGCTCACCTGCGAG-----GA 2400  
2363 TGGCAGTGCCTCAGAGAACTCCGTTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACAGCGCCCTGAGGAGTGGCGTGTCTATCGAGCAGAGCTCTCTCTGATGAGGCCAG 2460  
2423 CAGCAGCCTGTTTGAATCTTCTGCTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482  
2461 CAGTGGCTCTTTCAGAGCTTCTTCTAGCTTCTATGTGCCACACGCGCTGCGCATCAGGGG 2520  
2483 CAGGTGCTATACAGTGCAGGCGATCCCCCAGGCTCCAGGCTATCCACCTCTCTCTG 2542  
2521 CAGTCTCTAGTCCAGTGCAGGAGTCCCGAGGCTCCATCTCTCTCAGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAACAAAGTGTGTTGCTGAGGTGCGAGCGGATTTGGGTT 2602  
2581 CAGCTGTGCTACGCGACATGAGAACAAAGTGTGTTGCGGAGATTTCGCGGAGCGGCT 2640  
2603 GCTTTTACGTTTGTGATGACTTCTGTTGCTGAGCGCTCTACTTGGACCAAGCAAAAC 2662  
2641 GCTCCTGCGTTTGGTGGATGATTTCTTGTGTTGAGACCTCACCTCACCCAGCGGAAAC 2700



2663 CTTCCTCAGCACCTGTCCTCCATGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2722  
2701 CTTCCTCAGCACCTGTCCTCCATGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2760  
2723 GACAGTGGTGAATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2782  
2761 GACAGTGGTGAATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2820  
2783 GCTGCTCAGTGGTTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2842  
2821 GCGGCGCCAGCGGCTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2880  
2843 GTTCTGTGACTTCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2902  
2881 GCAGAGCGACTTCTCCAGCTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2940  
2903 TGCTCTCAAGCTGGAGAGCTGTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 2962  
2941 CGGCTTCAAGCTGGAGAGCTGTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3000  
2963 TCAGGCTTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3022  
3001 TCAGGCTTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3060  
3023 CAGATCTTCTGCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3082  
3061 CAGATCTTCTGCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3120  
3083 CCAGGCTTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3142  
3121 TCAGGCTTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3180  
3143 CTGCTATGCTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3193  
3181 CTGCTATGCTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3240  
3194 TGGCTCTTCTGCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3253  
3241 CGGCTTCTGCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3300  
3254 GGTGCTCTATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3313  
3301 GACTCGACACCGTGCATTCCTGAGTATTCCTCTGAGAGCTGCTGTCCTGAGTATGGTGCATGATAACTTCGAGAA 3360  
3314 GCTGTCGCGAAGCTTCGAGAGCTGCTGTCCTGAGTATTCCTGAGTATGGTGCATGATAACTTCGAGAA 3373  
3361 GCTGAGTCGGAAGCTTCGAGAGCTGCTGTCCTGAGTATTCCTGAGTATGGTGCATGATAACTTCGAGAA 3420  
3374 CCTAAGCAGACACTTTCAGACCACTTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTTCAGACCACTTTTGGACTAA 3454

## RESULT 10

Sequence 1, Application US/10044692

Publication No. US20030096344A1

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/044,692  
FILING DATE: 11-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/912,951  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
/note= "human telomerase reverse  
transcriptase (hTERT) catalytic protein  
component"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-044-692-1

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;  
QY 23 CCCGCCCTTGAGCACAATGACCCCGCTCCCTCGTTGGCCCGCGTGGCTCTCTGCTGCG 82  
Db 40 CCGGCCACCCCGGATGCGCGCGCTCCCGCTCGGAGCGCTGCGCTCCCTGCTGCG 99  
QY 83 CAGCGATACCGGAGGTGTGGCGCTGCGCAACTTTGTGCGCGCGCTGGGGCCGAGGG 142  
Db 100 CAGCCACTACCGGAGGTGTGGCGCTGCGCAACTTTGTGCGCGCGCTGGGGCCGAGGG 159  
QY 143 CAGCGCGCTTGTGCAACCCCGGAGCGCGCAAGATCTACCGCACTTTGGTTGGCCAAATGCCT 202  
Db 160 CTGGCGGTGTGTGACGCGCGGGGACCCCGCGGCTTTTCGCGCGCTGTGTGGCCAGTGCCT 219  
QY 203 AGTGTGATGCACTGGGCTCAGAGCTTCCACCTTCGCGCACTTTCTTCCTCCACAGGTGTC 262  
Db 220 GGTGTGCGTGCCTGGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 279  
QY 263 ATCCCTGAAAGAGCTGTGGCGAGGTTGTGCAAGACTCTGGAGCGCAACGAGAGAA 322  
Db 280 CTGCTGAAGAGCTGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 339

323 CGTGTGCTTTTGGCTTTGAGCTTGTGCTTAACGAGGCCAGAGGGCGGGCTTCCCATGSCCTT 382  
340 CGTGTGCTGCTTGGCTTTGCGCTTGTGAGCGGGCCCGGGGGGGGGGGGGGGGGGGGGGGGG 399  
383 CACTAGTACGCTGGGTAGTACTTTGCCCAACACATGTTATTAGAACCTTGCCTGTACAGTGG 442  
400 CACACACAGCGTGGCAGCTACTTGCCTCCCAACACAGTGAACGACGCTGGGGGAGCGG 459  
443 TGCATGATGCTACTGTGAGCCGAGTGGGGAGCAGCTCTGCTGCTACTCTGTGTCACACA 502  
460 GCGGTGGGGGGTGTGCTGCGCCCGCTGGGCGACGACGTGCTGCTTCACTGTGTCACG 519  
503 CTGTGCTCTTATCTTCTGCTGCGCCCGCCAGCTGTGCTACAGCTGTGCTGGGCTCTCCCT 562  
520 CTGCGCGCTCTTTGTGTGTGTGCTTCCAGCTGTGCTTACAGCTGTGCGGGCGCGCT 579  
583 GTACCAAAATTTGTGCGCACCAAGATATCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTG 622  
580 GTACAGCTCGCGCTGCTCACTCAGGCGCGCGCCCGCCACAGCTAGTGGACCCCGAAG 639  
623 CGGACCCGTGGCGAGGAATTTCACTAACCCTTAGCTTCTTACAACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGATGC-----GAAAGCGCTGGAAACCATAG 672  
683 TGGCAGGAAGCAGCAACCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 742  
673 CGTCAAGGAGCGCGGGTCCCGCTGGGCTGCCAGCCCGGCTGGAGGAGGGCGGGGG 732  
743 TCTACCAAGTACAGTGTGCTTCACTAGAGGCGGAGTGTATCTGCTCCGAGAT 802  
733 CAGTCCAGCGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
803 GGAGGAGGAGCCCAACAGGAGGCTGCTACCAACCCCATCAGGCAAAATCATGGGTGCCAAG 862  
793 GGAGGAGGAGCGCGTGGGCGAGGGGTCTGGGCGCCACCGCGGAGGAGCGGTGGACCGAG 852  
863 TCCTGCTCGTCCCCCGAGGTGCTACT-----GAGAGAAAGATTTGTCTCTAA 913  
853 TGACCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912  
914 AGGAAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970  
913 GGGTGGCTCTCTGCGACGCGCACTCCACCCCATCGTGGGCGCGCAGCACACCGCGGG 972  
971 CTCACAT---CTGTGCTGTACACCGCGCCAAATGCTTTGCTGCTGCTGCTGCTGCTGCT 1027  
973 CCCCCCATCCCATCGCGGCCACACGCTCCCTGGGACACGCTTGTCCCCGGGTGTACGG 1032  
1028 TGAGACAGACATTTCTTTACTCCAGGGAGATGGCCAAAGAGGCTCTAAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTTCTACTTCTCAGGCA---CAAGAGAGAGTGTGGGCGCTTCT 1089  
1088 CTTACTAGCAACCTTCAGCTAACTTGACTGGGGCGAGGAGTGTGCTGAGATCATCTT 1147  
1090 CTTACTAGCTCTCTGAGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149  
1148 TGTGGCTCAGGCGCTAGGACATCAGGACCACTCTGAGGACCAACCGCTATCGGCTG 1207  
1150 TGTGGTTCAGGCGCTGATGCGAGGAGTCCCGGAGGTTGCGGCGCTGCGGCGCTGCGG 1209  
1208 ATACTGGCAGATGCGGCGCTGTTCCAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1267  
1210 TCTACTGGCAAAATCGCGGCGCTGTTCTGAGCTGCTTGGGAAACCAAGCGAGTGGCCCTA 1269  
1268 TGTGAGATCTCTCAGTCACTGAGGCTTGGAAACAGCAAAACCAAGAGTGAAGATGC 1327  
1270 CGGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCCAGCAGCGGTGT 1329  
1328 CT-----TGAAACACGAGCC 1342  
1330 CTGTGCGCGGAGAGCCCGAGGCTCTGTGCGGCGCCCGGAGGAGGAGGAGGAGGAGGAG 1389  
1343 ACCGACCTCATGGATTTGCTCCCGCTGCAACAGCAGTCCCTGCGAGGATATGGTTTCT 1402

1390 CCGTGCCTTGGTGCAGCTGCTCGGCAGACAGAGCCCTTGGCAGGTGTACGGCTTCT 1449  
1403 TCGGGCGCTGCTCTGCAAGGTGTGTCTAGTCTCTGGGTACAGGCAAAATGAGCG 1462  
1450 GCGGGCTGCTCGCGCGCTGTGTGCCCGCAGGCTCTGGGCTCCAGGCAACAGAACG 1509  
1463 CCGCTTCTTTAAGAACTTAAAGATTCATCTCTGTGGGAAATACGGCAAGCTACT 1522  
1510 CCGCTTCTCTCAGGAACCAAGAAAGTTTCTCTCCCTGGGAAAGCATGCCAAGCTCT 1569  
1523 GCAGGAACTGATGTGAAAGATGAAGATAGGATTCACCTGCTCCCGCAGCAGCCCGGG 1582  
1570 GCAGGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGCTGCGCAGGAGCCAGG 1629  
1583 GAAGGACCGTGTCCCGCTGCGAGAGCAGCTGTGAGGAGAGAGATCTTGGCTACTTCT 1642  
1630 GGTGTGCTGTGTTCCGCGCCGAGCAGCAGCTCTGCTGAGGAGATCTTGGCCAGTCT 1689  
1643 GTTCTGCTGATGGACACATACGTGTGACAGCTGTAGTTCATTTTATCATCACA 1702  
1690 GCATGGCTGATGAGTGTGTACGTGTGAGCTGTCTGAGTCTTTTATGTACCGA 1749  
1703 GAGCACATTCAGAGAAACAGGCTCTTTCTTACCGTAAGATGTGTGAGAGCAAGTGA 1762  
1750 GACCACTTTCAAAGAAACAGGCTCTTTTCTTACCGGAAGATGTCTGAGCAAGTTGA 1809  
1763 GAGCATTTGAGTCAGGCAACACCTTGAGAGAGTGGGCTACGGGAGCTGTCAAGAGGA 1822  
1810 AAGCATTTGAAATCAGACAGCATTTGAAGAGGTGAGCTGCGGGAGCTGTGGAAGAGA 1869  
1823 GGTGAGCATCACAGGACACCTTGGCTAGCCATGCCATCTGCAGACTGCGTTTCATCCC 1882  
1870 GGTGAGGACATCGGGAAGCAGGCGCCGCTGTGAGCTGCAGACTCCGCTTTCATCCC 1929  
1883 CAAGCCAAACGGGCTCGGCGCCCATTTGTGAACATGATTTAGCATGGGTACAGAGCTTT 1942  
1930 CAAGCTGTAGCGGCTCGGCGGANTGTGAACATGAGACTACGCTGTTGGGAGCCAGAACGT 1989  
1943 GGGCAGAAAGAACAGGCGCCAGCATTTTCCAGGAGCTCTCAAGACTCTCTTTCAGATGCT 2002  
1990 CCGCAGAGAAAGAGGCGCGAGCGCTCACCTCGAGGTTGAAGGCACGTGTTTCAGCGTCT 2049  
2003 CAATATGAGCGAGCAAAACATCTCACTTATGAGGTCTTCTGTACTGGGTATGAATGA 2062  
2050 CAATACAGCGCGCGCGGCGCCCGCTCTGCGCGCTCTGTGCTGGGCTGGAG 2109  
2063 CATCTACAGGACCTGGCGGCGCTTGTGCTGCGTGTGCTGCTGCGACCAAGACACCCAG 2122  
2110 TATCCACAGGCGCTGGGCGACCTTCTGCTGCTGCTGCTGCGGGCCAGGACCCCGCCTGA 2169  
2123 GATGTACTTTTAAAGCAGATGTGACCGGGCGCTATGATGCCATCCCGCAGGTTAAGCT 2182  
2170 GCTGTACTTTTCAAGTGGATGTGACGGGCGCTGACACCATCCCGCAGGACAGCT 2229  
2183 GGTGGAGGTTTGTCAATATATGATCAGGCACTCGAGAGCAGTACTGTATCCGCACTA 2242  
2230 CAGGAGTCTATCCCAAGCATCATCA---AACCCAGAACAGTACTGGTGGCTCGGTA 2286  
2243 TGCACTGCTCCGAGAGATAGCAAGGCCAAGTCAACAGTCTTTTAGAGAGCAGGTCA 2302  
2287 TGCGGTGGTCCAGAGCGCGCCCATGGGACGCTCGCAAGGCTTCAAGAGCCAGCTCTC 2346  
2303 CACCTCTCTGAGCTCCAGCCATACATGGGCGAGTCTTAAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTACAGAGCTTCCAGCGGTACATGCGAGTGTGTTGGCTCACCTGCGAG-----GA 2400  
2363 TGGCAGTGCATGAGGAATCCCGTGTCTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2401 GACCCAGCCGCTGAGGGATGCGCTGCTCATCGAGCAGAGCTCTCTCCCTGAATGAGGCCAG 2460  
2423 CAGCAGCGCTGTTGACTTCTTCTGCACTTCTGCGTCAAGTGTGCTGCTGCTGCTGCTGCT 2482

2461 CAGTGGCTCTTCAGAGCTTCTCCTACGCTTCACTGTCACCAACGCGTGGCATCAGGG 2520  
2483 CAGTGGCTTATACGAGTGCAGGAGCTCCCGCAGGCTCAGCTTATCACCCTGCTCTG 2542  
2521 CAAGTCTTACCTCAGTGCAGGAGGATCCCGCAGGCTCATCTCTCCACGCTGCTCTG 2580  
2543 CAGTCTGTGTTTCGAGACATGAGAAACAGCTGTTTCTGAGTGCAGCGGATGGTT 2602  
2581 CAGCTGTGTACGGGACATGGAGAAACAGCTGTTTTCGGGATTCGGCGGACGGCT 2640  
2603 GCTTTAGTGTGTTGATGATCTTCTGTTGGTGAACCTCCTTCTGGACACAGCAAAAC 2662  
2641 GCTCTGCTGTTTGTGATGATTTCTTGTGTTGATGATTTCTTGTGTTGATGATTTCT 2700  
2663 CTTCCTCAGCAACCTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2722  
2701 CTTCCTCAGCAACCTGTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
2723 GACAGTGTGTAACCTTCCCTGTGAGAGCTGTGTAACCTGTGAGTGTGAGTGTGAGTGT 2782  
2761 GACAGTGTGTAACCTTCCCTGTGAGAGCTGTGTAACCTGTGAGTGTGAGTGTGAGTGT 2820  
2783 GCCTGTCTCAGTGCCTGTTTCCCTGTGAGAGCTGTGTAACCTGTGAGTGTGAGTGTG 2842  
2821 GCGGCTCAGAGCTGTTTCCCTGTGAGAGCTGTGTAACCTGTGAGTGTGAGTGTGAGT 2880  
2843 GTTCTGTGACTACTCAGTGTATGCCAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 2902  
2881 GCAGAGGAGCTACTCCAGCTATGCCAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 2940  
2903 TGTCTTCAAGCTGGAGAGCTATGCCAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 2962  
2941 CGCTTCAAGCTGGAGAGCTATGCCAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 3000  
2963 TCAGGCTTCTTCTAGACTTGCAGTGTGAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 3022  
3001 TCAGAGCTGTTTCTGAGTGTGAGAGCTCAATTAAGACAGAGCTCAGCTTCCAGAG 3060  
3023 CAAGATCTTCTGCTCAGGCTTACAGGTTCCATGATGATGATGATGATGATGATGATGAT 3082  
3061 CAAGATCTTCTGCTCAGGCTTACAGGTTCCATGATGATGATGATGATGATGATGATGAT 3120  
3083 CCAGGCTGTAGGAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3142  
3121 TCAGCAAGTGTGAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3180  
3143 CTGCTATGCTATCTGAGTGTGAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3193  
3181 CTGCTATGCTATCTGAGTGTGAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3240  
3194 TGGCTCTCTTCTCCTCAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3253  
3241 CGGCTCTCTCCTCAGAGCTCAGATTTCTTCTGAGGATCATCTCCAGCAAGCATCTG 3300  
3254 GGCTGTCTATCTGATCTACAAATGTTCTCTGAGGATCATCTCCAGCAAGCATCTG 3313  
3301 GACTCGACACCGTGTCACTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
3314 GCTGTGCGGAGCTCCAGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3373  
3361 GCTGAGTGTGAGAGCTCCGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420  
3374 CCTAAGCAGACACTTTCAGACCACTTTTGGACTAA 3407  
3421 ACTGCCCTCAGACTTTCAGACCACTTTCAGACTGA 3454

SULT 11

-10-044-539-1

Sequence 1, Application US/10044539

Publication No. US2003010093A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOWERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044,539

FILING DATE: 11-Jan-2002

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTERT"

/note= "human telomerase reverse

transcriptase (hTERT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-044-539-1

Query Match 45.8%; Score 1600.8; DB 15; Length 4015;  
Best Local Similarity 68.9%; Pred No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCGCGCTTTGAGCAATGACCCCGCTCTCGTTGCCCGCGTGGCTCTCTGCTGCG 82  
| | | | |  
DB 40 CCGCGCAACCCCGCGATCGCGCGCTCCCGCTGCCAGCCGCTCCCTGCTGCG 99  
| | | | |  
QY 83 CAGCGATACCGGAGGTGTGGCGTGGCAACCTTTGTGGGCGCTGGGCGCCGAGGG 142  
| | | | |  
DB 100 CAGCCACTACCGGAGGTGTGCTCCCTGGCCACTTGTGCGCGCTGGGCGCCCGAGGG 159  
| | | | |

143 CAGCGGCTTGTGCAACCCGGGAGACCGAAGATCTACCGACTTTGGTTGGCCAAATGCT 202  
160 CTGGCGGCTGTGTGAGCGCGGGACCCGGGGCTTTCCGCGCTGGTGGCCAGTGCT 219  
203 AGTGTGATGACATGGGGCTCAGAGCTCCACCTGCGGACCTTTCTTCCACACAGGTGTC 262  
220 GGTGTGCTGCTTGGAGCGCAGCGCGCCCGCGCGCCCTCTTCCGCGAGGTGTC 279  
263 ATCCCTGAAGAGCTGTGGCCAGGGTGTGAGAGACTCTGCGAGCGCAACGAGAGAA 322  
280 CTGCTGAAGAGCTGTGCGCCGAGTGTGAGAGCTGTGCGAGCGCGCGGCGAAGAA 339  
323 CGTGTGCTTTTGGCTTGTAGCTGTCTTACGAGGCGCAGAGCGGCGCTCCACGCTT 382  
340 CGTGTGCTTGTGCTTGTGCTTGTGAGCGGCGCGCGCGCGCGCGCGCGCGCTT 399  
383 CACTAGTAGCGTGTAGCTACTTTGCCCAACACTGTATTGAGACCCCTGCTGTGAGTGG 442  
400 CACCACAGCGTGTGCGAGCTACTTGCCCAACACCGTGACCGCACTGCGGGGAGCGG 459  
443 TGCATGATGCTACTGTGTGAGCGAGTGGCGAGCACTGCTGCTACTGCTGTGCGACA 502  
460 GCGGTGGGGCTGCTGCTGCGCGCGGTGGCGAGCGAGTGTGCTGCTGCTGCTGCGACG 519  
503 CTGTGCTCTTTATCTTGTGTGCGCCCGAGCTGTGCTTACAGGTGTGTGGGTCTCCCT 562  
520 CTGCGCGCTCTTGTGTGTGCTGCCAGCTGCCAGCTACAGGTGTGCGGGCGCGCT 579  
563 GTACCAATTTGTGCCACACAGGATCTGGGCTCTGTGTGCTGCTGTGCTGTGCTGTGCT 622  
580 GTACAGCTGCGCGCTGCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 639  
623 CCGACCCGTGGCGAGGAATTTCACTAACCCTTAGGTTCTTACACAGATCAAGAGCAGTAG 682  
640 GCGTCTGGATGC-----GAGCGGCTGTGAAACATAG 672  
683 TCGCCAGGAACGACCGAACCCTTGGCTTGCATCTCGAGGTACAAAGAGCATCTGAG 742  
673 GGTGAGGAGCGCGGGGTCCCTTGGCGCTGCCAGCCCGCGGTGCGAGAGCGCGGGG 732  
743 TCTCACCAGTACAGTGTGCTTGTGCTGTAGAGGCGCAGATCTATCTTGTCCGAGGT 802  
733 CAGTGTGCGAGCGAAGTGTGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 792  
803 GAGAGGAGGACCCACAGGAGGTGTACCAACCCCATCAGGCAATCATGGTGTGCGCAAG 862  
793 GAGCGGAGCGCGCTTGGCGAGGGTCTTGGGCGCGCGCGCGCGCGCGCGCGCGCGAG 852  
863 TCCTGCTCGGTCCCGCGGTGCTACT-----GCAGAGAGATTTGTCTCTAA 913  
853 TGACGCTGTTTGTGTGTGTCACTTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 912  
914 AGGAAGGTGTGTGACCTGAGTCTCTC---TGGGTGTGCTGTGCTGTAAACAAAGCCGAG 970  
913 GGTGTGCTCTTGTGCGAGCGCGCACTCCACCCCATCCGTGGCGCGCGCGCGCGCGG 972  
971 CTCACAT---CTGTGTGTGACACCCCGCGAATGCTTTCAGTGTGCGCGCAATTTAT 1027  
973 CCCCCATCCACATCGCGGCGCACCATCTCTCTGGGACAGCGCTTGTCCCGCGGTGTACGC 1032  
1028 TGAGACAGACATTTCTTTTACTCCAGGGAGATGGCCAGAGCGCTATAACCCCTCAT 1087  
1033 CGAGACCAAGCACTTCTTCTACTCTCAGCGA---CAGGAGCACTGCGGCGCTCTT 1089  
1088 CTTACTCAGCAACTCTCAGCTTAACTTGTGCGGCGAGAGCTGGTGGAGATCATCTT 1147  
1090 CTTACTCAGCTCTCTGAGGCGCGCGCTGACTGTGCGGTGCGAGGCTGTGGAGACCATCTT 1149  
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCACTGTGCGAGACACCGTCTATCGGTGCG 1207  
1150 TCTGGGTTCCAGGCGCTGTGATGCCAGGACATCCCGCAGGTTGCCCGCGCTGCCCGAGCG 1209

QY 1208 ATACTGGCAGATGGCGCCCTCTTCCAAACAGCTGTGTGTGAACCATGCGAGGTGCCAATA 1267  
Db 1210 CTACTGGCAAAATCGCGCCCTCTTCTGTGAGTGTGTTGGAAACCAACGCGAGTGCCTTA 1269  
QY 1268 TGTACAGCTCCTCAGTACATTTGAGAGTGTGGAACAGCAAAACAAAGGTGACAGATGC 1327  
Db 1270 CGGGGTGCTCTCAGAGCGCACTGCGCGCTGGAGCTGGGTCAACCCAGCAGCGGTGT 1329  
QY 1328 CT-----TGAAACACAGGCC 1342  
Db 1330 CTGTGCGCGGAGAAACCCAGGGCTCTGTGGCGCCCGCGAGGAGGAGACACAGAGCC 1389  
QY 1343 ACCGACCTCATGGAATTTGCTCCGCTGCAAGAGTCCCTGGGAGGTATATGTTTCT 1402  
Db 1390 CCGTCCCTGTGTGAGTGTCTCCGCGAGCAAGAGCCCTTGGCAGGTGTACGGCTTCT 1449  
QY 1403 TCGGGCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACCGGCAACAATGAGCG 1462  
Db 1450 GCGGGCTGTGCTGCGCGGCTGTGCGCCAGCGCTCTGGGCTCCAGGCAACAAGAACG 1509  
QY 1463 CCGCTTCTTAAAGACTTAAAGATTTGATCTCTGTGGGAAATACGCAAGCTTACCT 1522  
Db 1510 CCGCTTCTCAGAGAACCAAGAGTTTCTCTCCCTGGGAGAGCATGCCAAGCTCTGCT 1569  
QY 1523 GCAGGAACCTGATGTGAGATGAAAGTAGAGATTGCCACTGGCTCCGAGCAGCCCGGG 1582  
Db 1570 GCAGGAGCTGACGTGAGATGAGCGTGGGACTGCGCTTGGCTGGCAGGAGCCAGG 1629  
QY 1583 GAAGGACCTGTCCCGCTGAGAGACCGTGTGAGGAGAGGATCTTGGTACGTTCT 1642  
Db 1630 GGTGTGCTGTGTCCCGCGCAGAGCACCGTGTGCTGAGAGATCTTGGCAAGTTCT 1689  
QY 1643 GTTCTGGCTGATGGACACATACGTGTGTAGCTGTGTAGGTCACTTTTATACATCAGAG 1702  
Db 1690 GCACTGGCTGATGAGTGTGTGCTGTGAGCTGTCTGCTTCTTTATGTCAGGA 1749  
QY 1703 GAGCACTTCCAGAGAACAGGCTCTTCTTCTACGTAAGAGTGTGTGAGAGAGTGTGCA 1762  
Db 1750 GACACGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTGTGAGAGTGTGCA 1809  
QY 1763 GACATTTGAGTGTGAGAGACCTTTGAGAGTGTGCGCTACGGAGTGTGTGAGAGGA 1822  
Db 1810 AAGCATTTGAATCAGACAGCACTTGAAGGGTGTGAGTGTGCGAGCTGTGCGAGGAG 1869  
QY 1823 GGTGAGGATCAACAGGACACCTGTGTAGCCATCTGTGAGACTGTGCGCTTCTATCCC 1882  
Db 1870 GGTGAGGAGCATCGGAAAGCCAGCGCCCTGTGAGCTCCAGACTCCGCTTCTATCCC 1929  
QY 1883 CAAGCCCAAGCGGCTGCGCCCATTTGTGAAACATGATGATGATGATGATGATGATGATG 1942  
Db 1930 CAAGCTGTGCGGCTGCGGCGCATTTGTGAAACATGATGATGATGATGATGATGATG 1989  
QY 1943 GGGCAGAGAGAGAGCGGCCAGCATTTTCCAGCGCTCTCAAGACTCTCTTTCAGACTGCT 2002  
Db 1990 CCGCAGAGAAAGAGGCGCGAGCGCTCTCACTCGAGGCTGAAGCACTGTTCAGCGTGT 2049  
QY 2003 CACTATGAGCGGACAAACATCTCTCACTTATGGGTCTTCTGTACTGGGTATGATGA 2062  
Db 2050 CAACTAGAGCGGCGCGCGCCCTCTGCGGCGCTCTGCGGCGCTCTGCTGTGGGCTGTG 2109  
QY 2063 CATCTACAGGACCTGCGGCGCTTGTGTGCTGTGCTGTGCTGTGAGACAGACACCCAG 2122  
Db 2110 TATCCACAGGCGCTGCGGCACTTGTGTGCTGTGCTGTGCGGCGCGAGGAGCCGCGCTGA 2169  
QY 2123 GATGTATTTTGTAAAGCAGATGTGACCGGGGCTATGATGCCATGCCCGGAGGTAGCT 2182  
Db 2170 GCTGTATTTTGTCAAGGTGATGTGACGGGCGCGTACGACACCATCCCGGAGCAGGCT 2229  
QY 2183 GGTGAGGTGTGTGCAATATCATCATCAGGACTGTGAGAGAGCAGTACTGTATCCGCCAGTA 2242  
Db 2230 CACGAGGTCTGCGCGAGCATCTCA---AAGCCAGAACAGTACTGCTGCTGCTGCTGCT 2286  
QY 2243 TGCAGTGTGCGGAGAGATAGCCAAAGGCCAAAGTCCCAAGTCTCTTTAGGAGAGAGTCA 2302

2287 TCCGCTGTCAGAGGCGGCCCATGGGACGTCGCHAGGCCCTTCAGAGCCACGCTCTC 2346  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTTAAGCATCTGAGGATTCAGA 2362  
2347 TACCTTTGACAGACCTCCAGCCGCTACATGCGACAGTTCGTGGCTCACCTGCAG-----GA 2400  
2363 TCCAGTGCACTGAGGAACTCGTTGTCATCGACGAGCATCTCTATGATGAGAGCAG 2422  
2401 GACCAAGCCGCTGAGGGATGCGCTGTCATCGAGAGAGCTCCCTCCCTGATGAGGCCAG 2460  
2423 CAGCAGCCTGTTGACTTCTTCTCTGACATTCCTCGTCCACAGTCTCGTAAGATTTGTTGA 2482  
2461 CAGTGCCCTCTTCGAGCTCTTCTTCACTGCTTATGTCACCAACGCGTGGCATCAGGG 2520  
2483 CAGGTGCTATACGACGTCAGGAGCATCCCCAGGCTCCAGGCTTCCACCTGCTCTG 2542  
2521 CAACTCTACGTCCTGAGTCCAGGGATCCGAGGGCTCCATCTCTCCAGCTGCTCTG 2580  
2543 CAGTCTGTTTCGGAGACATGGAGAAACAAGCTGTTGCTGAGGTGAGCGGATGGGTT 2602  
2581 CAGCCTGTGCTACGCGACATGAGAAACAAGCTGTTGCGGGATTCGGCGGACGGCT 2640  
2603 GCTTTTACGTTTGTGATGATCTTCTGTTGAGCGCTCACTTGGACCAAGCAAAAC 2662  
2641 GCTCCTGCGTTTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCACGCAAAAC 2700  
2663 CTTCTCTCAGCACCTTGGTCCATGGGCTTCTGAGTATGGGTGACATGATAAATTGCAGAA 2722  
2701 CTTCTCTCAGGACCTTGGTCCGAGGTGTCCTGAGTATGGCTGCTGCTGAGTTCGGAA 2760  
2723 GACAGTGTGAACTTCCCTGTGAGACCTGGTACCTGCTGGTGGTGCAGCTCCATCAGCT 2782  
2761 GACAGTGTGAACTTCCCTGTGAGAAAGAGAGGCGCTGGGTGGCACGCTTTTGTTCAGAT 2820  
2783 GCTGCTCACTGCTGTTTCCCTGTGTTGGTGTGCTGCTGAGACTCAGACTTTGAGGT 2842  
2821 GCGGCGCACGCGCTATTCCCTGTGTCGGCTGTGCTGTGATACCGGACCTTGGAGGT 2880  
2843 GTTCTGTGACTACTCAGGTTATGCCAGACCTCAATTAAGACGAGCTCACCTTCACAG 2902  
2881 GCAGAGCGACTACTCCAGTATGCCCGACCTCCATCAGAGCCAGTCTCACCTTCAACCG 2940  
2903 TGTCTTAAGCTGGGAGACCAATCGGAAACAGCTCTGCTGCTGCTGTTGGTGTGAGTG 2962  
2941 CGGCTTCAAGGCTGGGAGGAAATGCGTCGCAAACTCTTTGGGCTCTTGCGGCTGAAGTG 3000  
2963 TCAGGCTCTATTCTAGACTTGCAGGTGAACAGGCTCCAGACGCTCTGCATCAATATATA 3022  
3001 TCACAGCTGTTTCTGGATTTGACAGTGAACAGCTCCAGACGCTGTGCACCAACATCTA 3060  
3023 CAGATCTTCTGCTTACGCTCAGGCTCAGGTTCAATGATGTGATTCAGCTTCCCTTTGA 3082  
3061 CAAATCTCTCTGCTGCAAGGCTACAGGTTTCAAGCATGTGTGCTGAGCTCCCATTTCA 3120  
3083 CCAGCGTGTAGGAAGAACCTCACATCTTCTGCGCATCATCTCCAGCAAGCATCCTG 3142  
3121 TCAGCAGGTTTGGAGAACCCCAATTTTCTGCGCTCATCTGACAGGCTCCCT 3180  
3143 CTGCTATGCTATCTGAAGGTCAAGAAATCCAGGAATGACATAAG-----GCCTC 3193  
3181 CTGCTACTCCATCTGAAGCCCAAGAACGCGAGGATGTGCTGGGGGCCAAGGCGCGCG 3240  
3194 TGGCTCTTCTCTGAGCGCGCATTTGGCTCTGCTPACCGAGCTTCTGCTCAGCT 3253  
3241 CGGCGCTCTGCGCTCGAGGCGCTGAGTGGTGTGCCACCAAGCNTTCTGCTCAAGCT 3300  
3254 GGTGCTCATTTCTGCTATCAATAATGTCTCTGGGACCTCTGAGGACAGCCCAAAACT 3313  
3301 GACTGACACCGCTGTCACTAGTCCCACTCTGCGGCTCACTCAGGACAGCCAGCGCA 3360  
3314 GCTGTGCGGAGCTCCAGAGCGCAATGACCATCTTAAAGCTGAGCTGACCCAGC 3373

Db 3361 CTGAGTCGAGAGCTCCGCGGAGCAGCTGACTGCTGAGGCGCGAGCCNACCCGGC 3420  
QY 3374 CCTAAGCACAGACTTTCAGACCAATTTTGGACTAA 3407  
Db 3421 ACTGCCCTCAGACTTCAAGACCATCTCTGGACTGA 3454  
RESULT 12  
US-10-385-882-1  
; Sequence 1, Application US/10385882  
; Publication No. US20030232409A1  
; GENERAL INFORMATION:  
; APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA  
; APPLICANT: FARRIS, James  
; APPLICANT: FOSTER, Douglas  
; APPLICANT: O'GRADY, Scott  
; TITLE OF INVENTION: IMMORTAL PORCINE CELLS  
; FILE REFERENCE: 110.01700101  
; CURRENT APPLICATION NUMBER: US/10/385,882  
; PRIOR FILING DATE: 2003-03-11  
; PRIOR FILING DATE: 2002-03-11  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1  
; LENGTH: 4027  
; TYPE: DNA  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: Nucleotide sequence encoding telomerase reverse transcriptase  
US-10-385-882-1

Query Match 45.8%; Score 1600.8; DB 12; Length 4027;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2379; Conservative 0; Mismatches 967; Indels 108; Gaps 9;

QY 23 CCCGCCCTTGAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGCGCTCTCTGCTGCG 82  
Db 41 CCCGCCCACCCCGCATGCGCGCGCTCCCGCTCCGAGCGCTGCTCTCTGCTGCG 100  
QY 83 CAGCGCATACCGGAGGTGTGGCGCTGGCAACTTTGTGCGCGCTGCGGCCCGAGG 142  
Db 101 CAGCCACTACCGCAGGTGTCTGCGCTGGCACTGTTGTGCGGCGCTGCGGCCCGAGG 160  
QY 143 CAGCGGCTTGTGCAACCCGCGGACCCGAAAGATCTACCGCACTTTGTTGCCCAATGCT 202  
Db 161 CTGCGCGCTGTGTGAGCGCGGAGACCCGCGGCTTTCCGCGCTGTGTGCGCCAGTCT 220  
QY 203 AGTGTGATGACATGGGCTTCAGACCTCCACCTGCGGACCTTTCTTCCACAGGTGTC 262  
Db 221 GGTGTGCTGCTTGGAGACGACGCGCGCCCGCCCGCGCGCTCTTCCGCGAGGTGTC 280  
QY 263 ATCCCTGAAGAGCTGTGCGCAGGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAA 322  
Db 281 CTGCTGAAGAGCTGTGCGCGCGAGTGTGTGAGAGGCTGTGTGAGCGCGGCGCGAAGAA 340  
QY 323 GGTGCTGCTTTTGGCTTTGAGCTTTAAGAGGCGAGGCGGCGCTCCATGCGCTT 382  
Db 341 GGTGCTGCTTTTGGCTTTGAGCTTTGAGAGCGGCGCGGCGCGCGCGCGCGCGCTT 400  
QY 383 CACTAGTAGCTGCTGAGTACTTTGCCCAACACTGTTATTGAGACCTCTGCTGTCAAGTG 442  
Db 401 CACCACAGCGTGCAGCTTACCTGCCCAACACGCTGACCGACGCACTGCGGGGAGCGG 460  
QY 443 TGCATGATGCTACTTGTGAGCGAGTGGGCGAGCAGCTGCTGCTACTTCTGCTGACCA 502  
Db 461 GCGTGGGGGCTGCTGTGTGCGCGCGTGGGCGAGCAGCTGCTGGTTCACTGCTGCGACG 520  
QY 503 CTGTGCTCTTTATCTTCTGCTGCGCCCGAGCTGTGCTACAGGCTGTGGGTCTCCCCCT 562  
Db 521 CTGCGGCTCTTTGCTGCTGCTGCTCCAGCTCCAGCTGCGGCTACAGGCTGTGGGCGCGCT 580  
QY 563 GTACCAAAATTTGTGCCACCAAGGATATCTGCGCCCTCTGTGCTGCTAGTTTACAGGCCAC 622

581 GTACAGCTCGGCGCTGCGCTACCTCAGCGCCGGGCCCGCCACACGCTAGTGGACCCGAAG 640  
623 CCGACCCGCTGGGAGGAATTTCACTAACCTTAGGTTCTTACACAGATCAAGAGCAGTAG 682  
641 GCGTCTGGGATG-----GAACGGGCTGGAACCATAG 673  
683 TCGCAGGAAGACCGGAACCCCTGCGCTTGCCATCTCGAGGTACAAAGAGCATCTCGAG 742  
674 CGTCAGGAGGCGCGGGTCCCTCGGCTGCGCAGCCCGGGTGGAGAGGCGCGGGG 733  
743 TCTCACCAGTACAAAGTGTGCTTCACTAAGAGGCGAGATCTATCTCTGCTCCGAGAGT 802  
734 CAGTGCACGCGAAGTCTGCGTGTGCCAAGAGGCCAGCGCTGGCGCTGCCCCGTGAGCC 793  
803 GGAGAGGAGCCCGACAGGAGGTGTACCAACCCCATCAGGCAAAATCATGGTGCCCAAG 862  
794 GGAGCGAGCCCGGCTTGGGAGGGGCTCTGGGCCACCCCGGCGAGGACGCGTGGACCGAG 853  
863 TCCTGCTCGGTCCCGAGGTGCTACT-----GCAGAGAAGATTTGTCTTCTAA 913  
854 TGACCGTGGTTTCTGTGTGTGTGTAAGTGCAGAGCCCGCGAAGAGCCACCTCTTTGGA 913  
914 AGGAAGGTGTGACCTGAGTCTCTC---TGGGTGCGTGTCTGTAAACAAAGCCCGAG 970  
914 GGGTGCCTCTCTGGACCGCGCCACTCCACCCCATCCGTGGCGCGCAGCACCCAGCGGG 973  
971 CTCCACAT---CTCTGCTGTCACACCCCGCCCAAAATGCTTTCAGCTCAGGCCATTTAT 1027  
974 CCCCCCATCCACATCGCGGCCACACAGTCTCTGGGACACGCTTGTCCCGCGGTGACGC 1033  
1028 TGAGACAGACATTTCTTTACTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCAT 1087  
1034 CGAGACCAAGCACTTCTCTACTCTCTCAGGCGA---CAAGGAGCAGCTCGCGCTCTCT 1090  
1088 CTTACTCAGAACCTCAGCCTTAATTGACTGGGCGCAGGAGACTGTGGAGATCATCTT 1147  
1091 CTTACTCAGTCTCTGAGGCGCCAGCCTGACTGCGCTCGGAGGCTCGTGGAGACCATCT 1150  
1148 TCTGGGCTCAAGGCTTAGGACATCAGGACCATCTCTGAGGACACACCGTCTATCGCTCG 1207  
1151 TCTGGTTCCAGGCCCTTGGATCGCAGGACTCCCGCAGGTTGCCCGGCTGCCCGCAGG 1210  
1208 ATACTGAGATCGGCGCTGTTCCAAAGTGTCTGTTGTAACCAATGACAGATGCCAATA 1267  
1211 CTACTGCAAAATCGGCGCTGTTCTGAGCTGTCTTGGGAAACCAACGCGCAGTGGCCCTA 1270  
1268 TGTGAGACTCTCAGGTACATTTGCGAGTTTCGAAACAGCAAAACACAGGTGACAGATGC 1327  
1271 CGGGGTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTACCCCAAGCAGCGGTGT 1330  
1328 CT-----TGAAACACAGGCC 1342  
1331 CTGTGCCGGGAGAAGCCCGAGGCTCTGTGGCGGCCCGAGGAGGAGACACAGACCC 1390  
1343 ACCGCACTCATGATTTGCTCGCTGCGCTGACAGCAGTCTCTGCGAGATATGTTTTCT 1402  
1391 CCGTGGCTGGTGAGTGTCTCGCGCAGACACAGCAGCCCTGGCAGGTGACGGTTCTGT 1450  
1403 TCGGGCTGTCTCTGCAAGTGTGTCTGTAGTCTCTGGGTACCCAGGCAAAATGAGCG 1462  
1451 GCGGGCTGCTGCGCGCTGTTGCCCGCCAGGCTCTGGGCTCCAGGCAACAGAACG 1510  
1463 CCGTTCTTTAAGAACTTAAGAGTTTCACTCGTTGGGGAATACGGCAAGCTATCACT 1522  
1511 CCGTTCTCAGGAACCAAGAAAGTTTCACTCTCCCTGGGGAAGCATGCAAGCTCTCGCT 1570  
1523 GCAGAACTGATGTGGAAGATGAAGTAGAGGATGCGCATGCGCTCCGAGCGCCCGG 1582  
1571 GCAGAGCTGACGTGGAAGATGAGCTGCGGGCTGCGCTTGGCTGCGAGAGCCAGG 1630  
1583 GAAGGACCGTGTCCCGCTCGAGACACCGTCTGAGGAGAGGATCTTGGCTACGTTCTCT 1642  
1631 GGTGGCTGTGTTCGGCGCGCAGAGCACCGTCTCGGTGAGGAGATCTTGGCCAAGTTCTCT 1690  
1643 GTTCTGGCTGATGAGACACATACGTGGTACAGTCTCTTAGGTCAATTTTACATCAAGA 1702  
1691 GCATGGCTGATGAGTGTGTAGCTGCTGAGTCTCTCAGGTCTTTTATGTACCGGA 1750  
1703 GAGCACATTCAGAGAAACAGCTCTTCTTACCGTAAGAGTGTGTGGAGCAAGTGA 1762  
1751 GACCACTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGCAAGTTGCA 1810  
1763 GAGCATTTGAGTTCAGGCAACACCTTGAGAGAGTGGCTACCGGAGCTGTCAAGAGGA 1822  
1811 AAGCATTTGAATCAGACAGCACTTGAAGAGGTGACCTGCGGAGCTGTTCGGAAGCAGA 1870  
1823 GGTGAGGATACACAGGACACTGCTGCTAGCCATGCCCATCTGAGAGATGCGCTTTCATCCC 1882  
1871 GGTGAGGAGACTCGGGAAGCCAGGCCCGCTCTGACGTCCAGACTCCGCTTTCATCCC 1930  
1883 CAAGCCCAACGGCTTCGGGCCCATTTGTGAACATGAGTTATAGCATGGGTACACAGGCTTT 1942  
1931 CAAGCTTGACGGCTTCGGCGCGATTGTGAACATGACATCGTCTGCGGAGCCAGAACGTT 1990  
1943 GGGCAGAAAGAGAGCGCCAGCAATTTCAACAGCGTCTCAAGACTCTCTTTCAGCATGCT 2002  
1991 CCGCAGAGAAAGAGGGCGCGAGCGTCTCACCTCGAGGCTGAAGGCACTGTTTCAGCGTCT 2050  
2003 CAACATGAGCGGACAAACATCTCTCACCTTATGGGTCTTCTGCTACTGGGTATGAATGA 2062  
2051 CAACCTACAGAGCGGCGCGCGCCCGGCTCTCTGGGCGCTCTGTGTCTGGGCTTGGACGA 2110  
2063 CATCTACAGGACTTCGGCGGCTTTGTGCTCGTGTGCGTCTCTGGAACAGACACCCAG 2122  
2111 TATCCACAGGGCTTCGGCGACCTTCGTGCTGCTGTGCGGCGCCAGGACCCGCGCTGA 2170  
2123 GATGTACTTTGTAGGACAGATGTGACCGGGCGCTATGATGCCATCCCGCAGGTAAGCT 2182  
2171 GGTGTACTTTGTCAAGGTGGATGTGACGGGCGCTGACACCATCCCGCAGGACAGGCT 2230  
2183 GGTGAGGTGTGTGCAATATGATCAGGCACTCGAGAGCAGCTACTGTATCCCGCAGTA 2242  
2231 CACGGAGGTTCATCGCCAGCATCATCA---AACCCAGAACACAGTACTGCGTGCCTGTA 2287  
2243 TGGAGTGTTCGGAGAGATAGCAAGGCCAAAGTCCACAGTCTCTTTAGGAGACAGGTAC 2302  
2288 TGGCGTGTGTCCAGAAAGGCGGCCCATGGGCAGCTCCGCAAGGCTTTCAAGAGCCAGCTC 2347  
2303 CACCTCTCTGACCTCCAGCCATACATGGGCGAGTCTCTTAAGCATCTGCAAGGATTCAGA 2362  
2348 TACCTTGACAGACCTCCAGCCGTATCATGCGACAGTTCTGTGGCTCACCTGCAG-----GA 2401  
2363 TGCCAGTGCATGAGAACTCCGTTGTATCGAGCAGAGCATCTCTATGAATGAGAGCAG 2422  
2402 GACCAAGCCGCTGAGGATGCGCTGTCATCGAGCAGAGTCTCTCCCTGATGAGGCCAG 2461  
2423 CAGCAGCTGTTTGTACTTCTTCGCTGCTCTCCGTCACAGTGTCTGTAAGATTGCTGA 2482  
2462 CAGTGGCTCTTCGACGCTCTTCTACGCTTCATGTGCCACACCGCGCTGGCATCAGGG 2521  
2483 CAGGTGCTATACGCAAGTGCAGGGCATCCCGCAGGGCTCCAGCCATCCACCTGCTCTG 2542  
2522 CAAGTCTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCTCTCCAGCTGCTCTG 2581  
2543 CAGTCTGTGTTTCGAGACATGGAGAACAGCTGTTTGTGAGGTGACGGGATGGGTT 2602  
2582 CAGCTGTGCTACGCGCAGCATGGAGAACAGCTGTTTGGCGGGATTCGGCGGGACGGGCT 2641  
2603 GCTTTTACGCTTTTGTGTGACTTCTGTGTGTGACGCTCACTTGGACCAAGCAAAAC 2662  
2642 GCTCTGCGTTTGTGTGATGATTTCTTGTGTGACACTCACCTCACCCACGCGAAGAAC 2701  
2663 CTTCTCTCAGCAACCTGCTGCTTCCATGGCGTTCTTGATATGGGTGATGATAAATTTGCAAG 2722  
2702 CTTCTCTCAGGACCTGCTGGTCCGAGGTGTCTCTGAGTATGCTGCTGCTGAGTCTTGGGAA 2761

2723	GACAGTGGTGAAC	TTCCTCGTGGAGC	CTGGTAGCCCTGGT	GAGCGGTGGT	GAGCTCCATACAG	CT 2782
2762	GACAGTGGTGAAC	TTCCTGTAGAGAG	CAGAGCCCTGGGT	GCGACGGCTTTGT	TCAGT 2821	
2783	GCTGTCTCACT	CCCTGTTTCCCT	GGTGGCTTGTCT	GCTGGACACTC	CAGACTTTGGAG	GT 2842
2822	GCGGGCCACG	CGCCTATTCCCT	CTGGTGGCTGTCT	GGATACCCG	ACCTGGAGGT 2881	
2843	GTTCGTGTACT	ACTCAGGTATG	CCCGAGACTCA	ATAAGACGAG	CGCTCACCTTCC	AG 2902
2882	CGACAGCACT	ACTCAAGTATG	CCCGACCTC	ATCAGAGCAG	CTCACCTTCA	ACC 2941
2903	TGTCCTTCAA	AGCTGGGAAG	CATTCGGGA	CAAGCTCT	CTCGGCTTGGG	TTGA 2962
2942	CGGCTTCAAG	GGCTGGGAG	GAACATGCG	CTCGAACTCT	TTTGGGCTTGG	CGCT 3001
2963	TCACGGTCTAT	TCTTAGACTT	GCAGGTGAAC	AGCCTCCAG	ACAGCTCGCAT	CAATATA 3022
3002	TCACAGCTG	TTTCTGGAT	TTTGCAGGT	GTGAACCTCC	AGCGGTGCG	CAACATCTA 3061
3023	CAAGATCTT	CTCTGCTCAGG	CTCAGGTTCC	ATGATGTGAT	TCAGCTTCCCT	TTGA 3082
3062	CAAGATCTCT	CTGCTGCAGG	CGTACAGTTT	TCAGCATGTG	CTGCAGCTCC	CAATTTC 3121
3083	CCAGCGTGT	TAGAGACCT	CACATTCCT	TTCTGGCATCAT	CTCCAGC	CAAGCATCTG 3142
3122	TCAGCAAG	TTTGGAAAC	CCCACTTTT	CTCGCGTCAT	CTCTGAC	CGGCTCCCT 3181
3143	CTGCTATGT	ATCTTGAAG	GTCAAGAT	TCAGGAAT	GCACATAAG	-----GCCTC 3193
3182	CTGTACTCT	CACTCTGAA	AGCAAGAA	CGCAGGAT	GTGCTGGGG	CGCAAGGCGCGC 3241
3194	TGCTCTCTT	TCTCTGAAG	CGGCACAT	CTGGCTCT	CTACAGGCTT	CTGTCTCAGCT 3253
3242	CGGCTCTG	CCCTCCGAG	GGCGTGCAGT	GGCTGTGC	CACCAAGCAT	TTCTGTCTCAGCT 3301
3254	GGCTGCTCA	TCTGTCAAT	GTCTCTT	GGGACCTCT	GAGGACAG	CGCCAAAAC 3313
3302	GACTCGACA	CCGTGTCACT	ACTGTGC	CACTCTCT	GGGTCAC	TAGGACAGCCAGCGCA 3361
3314	GCTGTCCG	GAAGCTCC	CAGAGGG	CAATGAC	CACTCTTA	AGCTGCAGCTGACCCAGC 3373
3362	GCTAGT	TCGGAAG	CTCCCGG	AGAGCTGT	ATGCTG	CCCTGGAGCCGCGACCAACCCGGC 3421
3374	CCTAAGCA	CAGACTT	TCAGAC	CACTTTT	GGACTAA 3407	
3422	ACTGGCCT	CTCAGACTT	CAAGAC	CACTCT	GTGACTCA 3455	

RESULT 13  
 -09-749-728B-32  
 Sequence 32, Application US/09749728B  
 Patent No. US20020142457A1  
 GENERAL INFORMATION:  
 APPLICANT: Umezawa, Akihiro  
 APPLICANT: Hata, Jun-Ichi  
 APPLICANT: Fukuda, Keiichi  
 APPLICANT: Ogawa, Satoshi  
 APPLICANT: Sakurada, Kazuhiro  
 APPLICANT: Gojo, Satoshi  
 APPLICANT: Yamada, Yoji  
 TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMYOCYTES  
 FILE REFERENCE: 00766.000043  
 CURRENT APPLICATION NUMBER: US/09/749,728B  
 CURRENT FILING DATE: 2001-09-17  
 PRIOR APPLICATION NUMBER: H11-372826  
 PRIOR FILING DATE: 1999-12-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-01148  
 PRIOR FILING DATE: 2000-02-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-07741  
 PRIOR FILING DATE: 2000-11-02  
 NUMBER OF SEQ ID NOS: 80

```
; SOFTWARE: PatentIn Ver.2.0
; SEQ ID NO 32
; LENGTH: 3396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; OTHER INFORMATION: (1)..(3399)
US-09-749-728B-32

Query Match          45.6%; Score 1594.6; DB 10; Length 3396;
Best Local Similarity 68.9%; Pred. No. 0;
Matches 2368; Conservative 0; Mismatches 959; Indels 108; Gaps

QY 39 ATGACCCGCGCTCTCGTTGCCCGCGGTGCGCTCTCTGCTGCGCAGCCGATACCGGAG 98
Db 1 ATGCGCGCGCTCCCCCGCTGCGAGCGTGGCTCTCTGCTGCGCAGCCACTACCGCGAG 60
QY 99 GTGTGGCGCGCTGGCAACCTTTGTGCGCGCGCTGGGGCCCGAGGCGCAGGCGCTTGTGCAA 158
Db 61 GTGTGCGCGCTGGCCACGTTGCTGCGCGCGCTGGGGCCCCAGGGCTGGCGGCTGGTGAG 120
QY 159 CCCGGGAGACCGAAGATCTACCGCACTTTGGTTGGCCCAATGCTTAGTGTGCAATGCTGG 218
Db 121 CGCGGGGACCCCGGGCGCTTTCCGCGCGCTGGTGGCCAGTGCCTGGTGTGCGTGCCTCG 180
QY 219 GGCTCACAAGCTCTCACGCTGCGCGACCTTTCTCTCCACCAGGTGCATCCCTGAAGAAGCTG 278
Db 181 GACGCACGCGCGCCCCCGCGCCCTCTTTCCGCAGAGTGTCTGCTCTGAAGAGACTG 240
QY 279 GTGGCCAGGGTTGTGCAGAGACTCTGCAGCGCAAGAGAGAAACGTGTGCGCTTTTGGC 338
Db 241 GTGGCGCGAGTGTCTGACAGAGCTGTGCGAGCGCGGGCGGAAGAACGTGTGCGCTTCGCG 300
QY 339 TTTGAGCTGCTTAACGAGGCGCAGAGCGCGCGCTCCCATGCGCTTCACTAGTAGCGTGCCT 398
Db 301 TTGCGGCTGTGTGACGGGCGCGCGGGGGCCCCCGAGGCGCTTCAACCACAGCGTGC 360
QY 399 AGCTACTTTGCCCAACACTGTATTATTAGAACCTCGGTGTCACTGCTGCATGGAATGCTACTG 458
Db 361 AGCTACTTGGCCCAACACGGTGTACCGACGCACTCGGGGGAGCGGGCGTGGGGGCTGCTG 420
QY 459 TTGAGCCAGTGGGGGAGGACCTGCTGCTCTACTGCTGSCACACTGTGCTCTTTATCTT 518
Db 421 CTGCGCCCGCTGGGCGACGAGTGTGTTTCACTGCTGTGCACGCTGCGCGCTCTTTGTG 480
QY 519 CTGTGTGCCCCAGCTGTGCTACAGGTGTGGGTGTCTCCCTGTACCAAAATTTGTGCG 578
Db 481 CTGTGTGCTTCCAGCTGCGCTTACAGGTGTGCGGGCGCGCTGTACACAGCTCGCGCT 540
QY 579 ACCAGGATATCTGGCCCTCTGTGTCCGCTAGTTTACAGGCCACCCGACCCGTGGGACGG 638
Db 541 GCCACTCAGGCCCGGCCCGGCCACACGCTAGTGTGACCCCGAAGGCGTCTGGGATGC 597
QY 639 AATTTCATCACTTAGTTGTTCTTACACAGATCAAGAGCAGTAGTGTGCGCAGAGACCGCG 698
Db 598 -----GAAAGGCGCTGGAAACATAGCGTCAAGGAGCGCGGG 633
QY 699 AAAAGCCCTGGGCTTTGCCATCTCGAGGATCAAGAGGCATCTGAGTCTCACCAAGTACAAGT 758
Db 634 GTCCCTCTGGGCTCTGCCAGCCCCGGGTGCGAGGAGCGCGGGGCGAGTGCACGCGAAGT 693
QY 759 GTGCTTTCAGCTAAGAGGCGAGATGCTATCTGTTCGAGAGTGTGAGAGGGAAGAGAGAG 818
Db 694 CTGCGGTTGGCCAAAGAGGCGCAGCGTGTGGCGCTGTGCCCTGTGAGCCGAGGCGAGCGCC 753
QY 819 AGGCAGGTGTCTACCAACCCCATCAGGCAAAATCATGTGTTGCCAAGTCTCTCTCGTCCCG 878
Db 754 GGGCAGGGGTCTCTGGGCGCCACCGGGCAGGACGCGTGGACCGAGTGACCGTGGTCTCTGT 813
QY 879 GAGGTGCGCTACT-----GCAGAGAAAGATTTGTCTTTAAAGGAAAGGTGTCTGAC 929
Db 814 GTGTGTGTCACTGTGCAGACCGCGCAGAGAGCAACCTCTTTGGAGGGTGCCTCTCTGCG 873
```



930 CTGAGTCTCTC---TGGGTGCGTGTGCTGTAAACACAAAGCCAGCTCCACAT---CTGTG 983  
> 874 ACGGCCACTCCACCATCCGTGGCGCGCAGACACCGCGGCGCCCCATCCACATCG 933  
> 984 CTGTACACACCCGCCAAAATGCTTTACAGCTCAGGCCATTTATTAGACACAGACATTTTC 1043  
> 934 CGGCCACCAAGTCCCTGGGACACGCCTTGTCCCGGTGTACGCCGAGACCAAGACATTC 993  
> 1044 CTTTACTCCAGGGAGAGTGGCCAAAGAGCGTCTAAACCCCTCAATTCCTACTAGCAACCTTC 1103  
> 994 CTCTACTCCTCAGCGCA---CAAGAGACAGCTCGGCGCCCTCTCTCTACTACAGCTCTCTG 1050  
> 1104 CAGCCTAACTTTGACTGGGGCCAGAGACTGGTGGAGATCATCTTTCTGGGCTCAAGGCT 1163  
> 1051 AGGCCACAGCTGACTGGCGTCTGGAGCTCTGGAGACCATCTTTCTGGGTTCAGAGGCC 1110  
> 1164 AGGACATCAGAACCATCTTCGAGGACACACCGCTCTATCGGTCTGATCTGCGAGATCGG 1223  
> 1111 TGGATGCCAGGACTCCCGCAGGTTGGCCCGCTGCCCCAGCGCTACTGGCAATGGG 1170  
> 1224 CCCGTGTTCCACAGCTGCTGTGAACCATGACAGTGGCAATATGTACAGCTCCTCAGG 1283  
> 1171 CCCCTGTTCTGGAGCTGCTTGGGAACAGCGCAGTGGCCCTTACGGGTGTCTCTCAAG 1230  
> 1284 TCACATTCAGAGTTTCGAACAGCAAAACCAACAGGTGACAGATGCCT----- 1329  
> 1231 ACGCACTGCCGCTGCAGCTGCGGTCAOCCAGCAGCGGTCTCTGTGCGCGGAGAG 1290  
> 1330 -----TGACACACAGCCCGGAGGAGGACACAGACCCCGTGCCTCATGTGAT 1358  
> 1291 CCCAGGCGCTGTGGCGGCGCCCGAGGAGGAGGACACAGACCCCGTGCCTGTGTGAG 1350  
> 1359 TTGCTCGCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCTTCGGACCTGTCTCTGC 1418  
> 1351 CTGCTCCGCGCAGCACAGACGCCCTTGCAGGTGTACGGCTTCTGGCGGCTCTGCTGCGC 1410  
> 1419 AAGTGGTGTCTGTCTGTCTGGGTAACAGGACAAATAGCGCGCGCTTCTTTAAGAAC 1478  
> 1411 CGGTGTGTGCCCCAGGCGCTTGGGGCTCCAGGCACAAAGCGCGCTTCTCTAGGAAC 1470  
> 1479 TTAAGAAGTTTCACTCTGTTGGGMAATACGGAAGTATCACTGCAGGAACGTATGTGG 1538  
> 1471 ACCAAGAAATTCACTCTCTGGGAGACATGCCAGCTCTCGTGCAGGAGCTGAGCTGG 1530  
> 1539 AAGATGAAGTAGAGATGTCCACTGTGCTCCGAGCAGCGCGGGAAGACCGTGTCCCC 1598  
> 1531 AAGATGAGCGTGGGACTGGCTTGGCTGGCAGGAGGCCAGCGGTTGGCTGTGTTCCG 1590  
> 1599 GCTCAGAGCACCGTCTGAGGAGAGGATCCTGGCTACGTTCTGTTCTGCGTGTGATGGAC 1658  
> 1591 GCGCAGAGCACCGTCTGCGTAGGAGATCCTGGCAAGTTCCTGCACCTGGCTGTGATGAGT 1650  
> 1659 ACATACGTGGTACAGCTGCTTAGTCAATCTTTTACATCACAGAGACATTCACGAAG 1718  
> 1651 GTGTACGTGCTGAGCTGCTCAGTCTTTCTTTTATGTACAGGACCAACAGTTTCAAAAG 1710  
> 1719 AACAGGCTCTTCTTACCGTAAGAGTGTGTGAGCAGAGCTGCAGAGCATTTGAGTACAG 1778  
> 1711 AACAGGCTCTTTTCTACCGGAAGAGTGTGTGAGCAAGTTCGAAAGCATTTGGAATCAGA 1770  
> 1779 CAACACCTTGAGAGTGGCGCTACGGGAGCTGTCAACAGAGAGGTGAGGATCACCAAG 1838  
> 1771 CAGCACTTGAGAGGGTGCAGCTCCGGAGACTGTCCGAGCAGAGGTGAGGAGCATTCGG 1830  
> 1839 GACACCTGGCTAGCCATGCCCATCTGTGAGACTGGCTTATCCCCAAGGCCCAACGCGCTG 1898  
> 1831 GAAGCCAGGCCCGCCCTGCTGACGCTCCAGACTCCGCTTCAATCCCCAAGCCTGACCGGCTG 1890  
> 1899 CGGCCCATTTGAACATGAGTTATAGCATGGTACAGAGCTTTTGGCAGAGAGGAGCAG 1958  
> 1891 CGGCCGATTTGAACATGAGTACGCTCGTGGGAGCCAGAACGTTCCGAGAGAAAGAGG 1950  
1959 GCCCAGCATTTCAACCAGCGTCTCAAGACTCTCTTTCAGCATGCTCAACTATGAGCGGACA 2018  
1951 GCCGAGCGTCTCACCCTCAGAGGGTGAAGCACTGTTCAGCGTCTCAACTAGACGGGCG 2010  
2019 AAACATCTCTCACTTATGGGGTCTTCTGTACTGGGTATGAATGACATCTACAGACCTGG 2078  
2011 CGGGGCCCGCGCTCTCTGGCGCTCTGTCTGGCGCTGGACGATATCCACAGGGCCTGG 2070  
2079 CGGGCCTTTGTCTGCTGCTGTGGTCTCTGACACGACAGACCCAGAGTGTACTTTGTTAAG 2138  
2071 CGCACCTTCGTCTCGTGTGGGGCCAGACCCGCGCTGAGCTGTACTTTGTCAAG 2130  
2139 GCAGATGTACACCGGCGCTTATGATGCCATCCCCAGAGGTAAAGCTGGTGGAGTGTGTGCC 2198  
2131 GTGGATGTACCGGGCGGTACGACACATCCCCAGAGACAGCTCACGGAGTCTATCGCC 2190  
2199 AATATGATCAGGCACTCGGAGAGCAGTACTGTATTCGCCAGATATGAGTGGTCCGGAGA 2258  
2191 AGCATCATCAAAACCCC---AGAAACAGTACTGCGTGGCTGCGTATGCGGTGGTCCAGAA 2247  
2259 GATAGCCAAAGGCCAAGTCCACAAGTCTTTTAGGAGACAGGTCCACCCCTCTCTGACCTC 2318  
2248 GCCGCCATGGGCACTCCGCAAGGCCCTTCAAGAGCCACGCTCTCTACCTTGACAGACCTC 2307  
2319 CAGCCATACATGGGCGAGTTCCTTAAGCATCTGCAGGATTCAGATGCCAGTGCACACTGAG 2378  
2308 CAGCGCTACATGCAGCAGTTCGTGGCTCACCTGCAG-----GAGACCGCGCTGAGG 2361  
2379 AACTCCGTTGTCTCGAGCAGAGCATCTATGATGAGAGCAGCAGCAGCGCTGTTTAC 2438  
2362 GATGCGCTGTCTCAGCAGAGCTCTCCCTGTAATGAGCCAGCAGTGGCTCTTTCGAC 2421  
2439 TTCTTCTCTCACTTCTCGCTCACAGTGTCTGAAGATTTGGTGAAGTGTATACGCGAG 2498  
2422 GTCTTCTCAAGTTCATGTGCGCACACCGCGTGGCATACAGGGCAAGTCTCTCGTCCAG 2481  
2499 TGCAGGGCATCCCCAGGGCTCCAGCCTATCCACCCTGCTCTGAGTCTGTGTTCCGA 2558  
2482 TGCAGGGGATCCCCAGGGCTCCATCTCTCCACGCTGCTCTGACGCTGTGTACGGC 2541  
2559 GACATGAGAAACAAGCTGTTTCTGAGGTGACGGGATGGGTGCTTTTACGTTTGT 2618  
2542 GACATGAGAAACAAGCTGTTTCTGCGGGATTCGGCGGAGCGGCTGCTCTCGGTTGGTG 2601  
2619 GATGAATTTCTGTGTGTGACGCTCACTTGTGAACGAAGCAAAAACCTTCTCAGCACCTG 2678  
2602 GATGAATTTCTGTGTGTGACACCTCACCTCACCTCACCGCGGAAAAACCTTCTCAGGACCTG 2661  
2679 GTCCATGGGTTCTCTGAGTATGGGTGCATGATAAATTTGCAGAGACAGTGGTGAACCTTC 2738  
2662 GTCCGAGGTGTCTCTGAGTATGCTGCGTGTGAACCTTTCGAGAGACAGTGGTGAACCTTC 2721  
2739 CTTGTGAGGCTGGTACCTCGGTGTGAGCTCCATACAGCTGCTGCTGCTCACTGCTG 2798  
2722 CTTGTAGAAGACAGAGCGCTGGGTGGCAGCGCTTTTGTTCAGATGCCGCGCCACGGCCTA 2781  
2799 TTTCCCTGCTGTGGCTTGTGCTGAGCAGCTCAGACTTTTGGAGGTGTTCTGTGACTACTCA 2858  
2782 TTCCCTGTGGCGGCTGCTGCTGTGATACCGGACCTTGGAGGTGCAGAGGACTACTTCC 2841  
2859 GGTATGTGCCAGACCTCAATTAAGACGAGCGCTCACCTTCCAGAGTGTCTTCAAAGCTGGG 2918  
2842 AGCTATGCCCGGACCTCCATCAGAGCGAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGG 2901  
2919 AAGACCATGGGAAACAGCTCCTGTGCTGCTTGGGTGAAGTGTACGGTCTATTCTTA 2978  
2902 AGAACTATCGTGGCAACTCTTTGGGGTCTTGGGGCTGAAGTGTCAAGGCTGTCTG 2961  
2979 GACTTCAGGTGAACAGCCTCCAGACAGTGTGCATCAATATATACAAGATCTTCTGCTT 3038  
2962 GATTTGAGGTGNAACAGCTCCAGAGGTGTGACCAACATCTACAGATCTCTCTGCTG 3021  
3039 CAGGCGCTACAGGTTCCATGCAATGTGTGATTTCAGTCTCCCTTTGACACGCGTGTAGGAAG 3098

3022 CAGGCGTACAGGTTTACGCGATGTGTGCTGAGCTCCCTCCATTTATCAGCAAGTTTGAAG 3081  
3099 AACCTCACATTTCTTCTGGGCAATCATCTCCAGCCAGCAATCTCTGCTATGCTATCTG 3158  
3082 AACCCACATTTTCTGCGCTCATCTCTGACAGGCTCCCTCTGCTACTCATCTG 3141  
3159 AAGGTCAAGATCCAGGAATGACACTAAG-----GCTCTGCTCTCTTCTCTCT 3209  
3142 AAAGCCAGAAACGAGGAGTGTCTGCTGGGGGCAAGGCGCGCGGCTCTGCTCTCC 3201  
3210 GAAGCGCACATTTGCTGTGTACCAAGCCCTTCTGCTGCTCAAGCTGGTGTCTCATTTGTC 3269  
3202 GAGGCGGTGCACTGGTGTGTCCACCAAGCAATTTCTGCTCAAGCTGACTCGACACCGGTGTC 3261  
3270 ATCTCAAAATGTCTCTGCGACCTCTGAGGACAGCCCAAAATCTGCTGTGCGGAGCTC 3229  
3262 ACTAGTGCCTACTCTGCGGTGCTCTGAGGACAGCCCAAGCTGAGTGTGGAAGCTC 3321  
3330 CCAGAGGCGCAATGACCATCTTAAAGCTGCACTGACCCAGCCCTAAGCAGACAGCTTT 3389  
3322 CCGGGAGACGCTGACTGCTGCTGAGGCGCGAGCCAAACCGGCACTGCCCTCAGACTTC 3381  
3390 CAGACCAATTTGGAC 3404  
3382 AAGACCATCTCTGGAC 3396

SUBT 14  
-10-205-629-1  
Sequence 1, Application US/10205629  
Publication No. US20030049236A1  
GENERAL INFORMATION:  
APPLICANT: Kassen, Thomas  
APPLICANT: Jensen, Thomas  
APPLICANT: Rattan, Suresh  
TITLE OF INVENTION: Immortalized Stem Cells  
FILE REFERENCE: 006148.00002  
CURRENT APPLICATION NUMBER: US/10/205,629  
PRIOR FILING DATE: 2002-07-26  
PRIOR APPLICATION NUMBER: 60/315939  
PRIOR FILING DATE: 2001-08-29  
PRIOR APPLICATION NUMBER: PA 2001 01148  
PRIOR FILING DATE: 2001-07-27  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 3453  
TYPE: DNA  
ORGANISM: Homo sapiens  
-10-205-629-1

Query Match 45.6%; Score 1593.6; DB 15; Length 3453;  
Best Local Similarity 68.9%; Pred. No. 0;  
Matches 2370; Conservative 0; Mismatches 964; Indels 108; Gaps 9;  
35 CACAATGACCCGCTCTCTGTCGCGCGTGGCTCTCTGTCGCGAGCGATACCG 94  
6 CACCAATGCGCGCTCTCTGTCGCGCGTGGCTCTCTGTCGCGAGCGACTACCG 65  
95 GAAGTGTGGCGCTGGCAACTTTGTGCGCGCTGGGCGCGGAGGCGAGCGGCTTGT 154  
66 CGAGTGTGCGCTGGCCACGTTGTCGCGCGCTGGGCGCGGAGGCGGCTGCTGT 125  
155 GCAACCCGGGACCCAGATCTACCGACTTTGGTGGCCCAATGCTAGTGTGATGCA 214  
126 GAGGCGGAGGACCCGCGGCTTTCCGCGCTGTGTCGCGAGTGTGTCGCTGTC 185  
215 CTGGGCTCAGAGCTCCACCTGCGGACTTTCTTCCACAGGTGTCATCTCCCTGAAAGA 274  
186 CTGGGACGACCGGCGCGCGGCTCTCTTCTGCGAGGTGTCCTGCTGCTGAAGA 245  
275 GCTGTGGCGAGGTTGTGACAGACTCTGCGAGCGCAACGAGAGAAACGCTGCTGCTTT 334

Db 246 GCTGTGCGCGAGTGTGACAGGCTGTGCGAGCGCGCGCGAAGAGCGTGTGCGCTT 305  
QY 335 TGGCTTTAGCTGCTTAACGAGGCGAGAGCGGCGCTCCCATGCGCTTCACTAGTAGCGT 394  
Db 306 CGGCTTGGCGTGTGAGACGGGCGCGCGGCGCGCGCGGCGCTTCAACACCGCGT 365  
QY 335 CGGTAGCTACTTGGCCCAACACTGTTATTGAGACCTGCTGCTGCTGCTGCTGCTGCT 454  
Db 366 GCGGAGCTACTCTGCGCCCAACACGCTGACCGAGCTGCGGGGAGCGGGGCGTGGGGGCT 425  
QY 455 ACTGTGAGCGAGTGGGCGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 514  
Db 426 GCTGTGCGCGCGTGGGCGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 485  
QY 515 TCTTCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574  
Db 486 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545  
QY 575 TGCCACACGAGTATCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634  
Db 546 GCTGCGCACTCAGCG 605  
QY 635 CAGGAATTTCACTAACTTAGTGTCTTAAACAGATCAAGAGCAGTAGTGTGCGGAGAGC 694  
Db 606 C-----GAAACGGGCTTGAAACCATAGCGTCAGGAGGCG 638  
QY 695 ACCGAAACCGCTTGGCTTGCATCTCGAGTACAAGAGGCACTCTGAGTCTCACCAGTAC 754  
Db 639 CGGGTCCCTTGGCGCTTGCAGCGCGCGGTGCGAGGCGCGGGGCGAGTGCACGCG 698  
QY 755 AAGTGTGCTTCACTAAGAGCGCGAGATGCTATCTGCTCCGAGAGTGGAGGAGGAGC 814  
Db 699 AAGTGTGCTTGGCG 758  
QY 815 CCAGGAGGCTGCTACCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874  
Db 759 GCTTGGCGAGGCTTCTGCG 818  
QY 875 CCGGAGTGTGCTACT-----GCAGAGAAAGATTTGCTTCTTAAAGGAAAGGTGTC 925  
Db 819 CTGTGTGCTGCTCACTGCGCAGACCGCGCGAGAGCGCACTCTTTGGAGGCTGCGCTCTC 878  
QY 926 TGACCTGAGTCTCTC---TGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979  
Db 879 TGGCAGCGCGCACTTCCCAACCATCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 938  
QY 980 TCTGCTGTCAACCGCGCGCAAAATGCTTTCAGCTCAGCGCAATTTATTGAGACGAGCA 1039  
Db 939 ATCGCGCGCAACGCTCTTGGGACACGCTTGTCTCCCGGCTGACCGCGAGACCAAGCA 998  
QY 1040 TTTCTTTTACTCCAGGAGATGCGCAAGAGCGTCTAAACCGCTCTATTCTTACTCAGCAA 1099  
Db 999 CTTCTCTTCTCTCTCAGGCGA---CAAGGAGCAGCTGCGCGCTCTTCTTACTCAGCTC 1055  
QY 1100 CTTCCAGCTTAACTTGACTGGGCGAGGAGCTGCTGAGAGATCATCTTTTGGGCTCAAG 1159  
Db 1056 TCTGAGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115  
QY 1160 GCCTAGGACATCAGGACCACTCTGCGAGGACACACCGCTCTATGCGCTGATAGTGGCAGAT 1219  
Db 1116 GCGCTGATGCCAGGACTCCCGCAGGTTGCCCGCTGCGCGCGCTACTTGGCAAT 1175  
QY 1220 GCGCGCGCTGTTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279  
Db 1176 GCGCGCGCTGTTCTGAGCTGCTTGGGAAACACGCGGAGTGGCGCTTACGGGCTGCTCT 1235  
QY 1280 CAGTGTCAATTTGAGGTTTTCGAAACCAAAACCAAGGTGACAGATGCTCT----- 1329  
Db 1236 CAGACGCACTGCGCGCTGCGAGCTGCGGTCAACCGAGCGAGCGGTGCTGTGTCGCGCGGA 1295  
QY 1330 -----TGAACACAGCG 1354

1296 GAAGCCCCAGGGCTCTGTGGGGCCCCGAGGAGGAGACACAGACCCCGCTGCTGGT 1355  
1355 GGAATTGCTCCGCTGCACAGAGTCCCTGGGAGGTATATGTTTTCTTCGGGCTGTCT 1414  
1356 GCAGCTGCTCCGCGACACAGAGCCCTGGCAGGTACGGCTTCGTGGGGCTGCT 1415  
1415 CTGCAAGGTGGTCTGCTAGTCTCTGGGGTACAGGCACATGAGCGCGCTCTTTAA 1474  
1416 GCCTGGCTGGTCCCGCCAGGCTCTGGGGCTCCAGGGACAAACGAAACCGCTTCCTCAG 1475  
1475 GAACTTAAAGAACTTCATCTCGTTGGGAAATACCGCAAGCTATCACTGCAGGAAGTAT 1534  
1476 GAACACCAAGAGTTTCACTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGAC 1535  
1535 GTGGAAGATGAAGTATAGAGATGACCATCTGGCTCCGACAGCCCGGGGAAGGACCGTGT 1594  
1536 GTGGAAGATGAGCGTCCGGACTGCGCTTGGCTGGCGAGGAGCCAGGGGTGGCTGTGT 1595  
1595 CCCGCTGCAGAGCACCGTCTGAGGGAGAGGATCCTGGTACGTTCTCTGCTGAT 1654  
1596 TCCGCGCGAGAGCACCGTCTGGTGAGGAGATCTGGCCAAAGTTCTTGCATGGCTGAT 1655  
1655 GGAACATACGTTGATAGCTGCTAGGTGATCTCTTTTACATCAAGAGAGACATTCGA 1714  
1656 GAGTGTGTACGCTGCTGAGCTGCTCAGGTCTTTCTTTTATGTACGGAGACACGTTTCA 1715  
1715 GAAGAACAGGCTCTTCTTCTACCGTAAAGAGTGTGGAGCAAGCTGCAGAGCATTTGAGT 1774  
1716 AAAGAACAGGCTCTTTTCTTACCGCGGAGTGTCTGGAGCAAGTTGCAAGCATTTGAT 1775  
1775 CAGGCAACACCTTTGAGAGATGCGGTACCGGAGTGTCAAGAGGAGGTGAGCATCA 1834  
1776 CAGACAGACTTGAAGAGGTGCAGCTCGGAGTGTTCGGAAGCAGAGGTGAGGAGCA 1835  
1835 CCAGGACACTGCTGAGTACCATGCCATCTGCAGCTGCGCTTATCCCGAGCCCAAGG 1894  
1836 TCAGGAAAGCAGGCGCGCTCTGCTGAGTCCAGACTCCGCTTATCCCGAGCCCTGAGG 1895  
1895 CCTCGGCGCCATTTGTAAATGAGTTATAGCATGGGTACAGAGCTTTGGGAGAGGAA 1954  
1896 GCTCGGCGCGATTTGTAAATGAGTTATAGCATGGGTACAGAGCTTTGGGAGAGGAA 1955  
1955 GCAGGCGCGATTTACCGAGGCTCTCAGACTCTCTTACGATGCTCTCAACTATGAGG 2014  
1956 GAGGCGCGAGCGTCTACCTCGAGGCTGAAGGCACTGTTCAAGCTGCTCACTACGAGCG 2015  
2015 GACAAACATCTTCACTTATGGGTCTTCTGATCGGTATGATGATCATCTACAGGAC 2074  
2016 GCGCGGCGCGCGCTCTCTGGGCGCTCTGCTGGGCTGGAGCATATCCACAGGCG 2075  
2075 CTGCGGCGCTTTGTGTGGTGTGGTGTCTGGAGCAGACACCGAGATGATCTTTGT 2134  
2076 CTGCGGCACTTCTGCTGGCTGTGGGCGGAGGAGCCCGCGCTGAGCTGTACTTTGT 2135  
2135 TAAGGAGATGTGACGGGCTATGATGATGATGATGATGATGATGATGATGATGATGAT 2194  
2136 CAGGTGATGTGACGGGCGGTACGACACATCTCCCGAGGACAGGCTCACGGAGGTAT 2195  
2195 TGCCAAATATGATAGGCACTCGGAGGACAGTATGATGATGATGATGATGATGATGATG 2254  
2196 CGCCAGCATATCAAAACCCG---AGAACAGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2252  
2255 GAGAGATAGCCAGCCAGTCCACAGTCTTTAGGAGCAGGTCACACCGCTCTCTGA 2314  
2253 GAAGCGCCCATGGGACGCTCGCAAGGCTTCAAGAGCAGCTCTTACCTTGAAGA 2312  
2315 CCTCAGCCATACATGGGCGGAGTTCTTTAAGCATCTGAGGATTCAGATGCCAGTGCAT 2374  
2313 CCTCAGCGGTACATGCGACAGTTCTGGTGGCTCAGTGCAG-----GAGACAGCCGCT 2366  
2375 GAGGAATCCGTTGTATCGAGGAGGAGCATCTCTAATGATGAGGAGGAGGAGGCTGTT 2434  
2367 GAGGATGCGGCTGCTATCGAGGAGGAGCTCTCTCTGATGAGGCGGAGGAGTGGCTCTT 2426

RESULT 15

US-10-105-616-1

; Sequence 1, Application US/10105616

2435 TGACTTCTTCTGCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2494  
2427 CGACGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2486  
2495 GCAGTGCAGAGGCTATCCCGCAGGGCTCCAGCTATCCAGCTGCTCTGCAAGTCTGCTGTT 2554  
2487 CAGTGCAGAGGATCCCGCAGGGCTCCATCTCTCCAGCTGCTCTGCAAGCTGCTGCTA 2546  
2555 CGGAGCATGAGGAACAAGCTGTTTCTGAGGTGCTGAGGTGCTGAGGTGCTGCTGCTGCTGCT 2614  
2547 CGGCGACATGAGGAACAAGCTGTTTGGGGGATTCGGCGGGAGCTGCTCTGCTGCTGCT 2606  
2615 TCTTCATCACTTCTGTTGGTGAAGCTCACTTGGACCAAGCAAAACCTTCTCTCAGCAC 2674  
2607 GGTGATGATTTCTTGTGTGACACCTCACTCACCACCGCAAAACCTTCTCTCAGCAC 2666  
2675 CTTGTCATGCGCTTCTGAGTATGGGTGATGATAAATTGCAAGAAGACAGTGTGAA 2734  
2667 CTTGTCAGAGTCTTCTGAGTATGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2726  
2735 CTTCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2794  
2727 CTTCTCTGAGGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2786  
2795 CTTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2854  
2787 CTTATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2846  
2855 CTTAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2914  
2847 CTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2906  
2915 TGGGAGAGCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2974  
2907 TGGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2966  
2975 TCTAGACTTGCAGGTGAACAGGCTTCCAGAGTCTGATGATGATGATGATGATGATGATGATGCT 3034  
2967 TCTGATTTGCAAGGTGAACAGGCTTCCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3026  
3035 GCTTCAAGGCTACAGTTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3094  
3027 GCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3086  
3095 GAAGAACCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3154  
3087 GAAGAACCTTCACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3146  
3155 CTTGAGGCTCAAGATCCAGGATGACACTAAAG-----GCCTTGGCTCTTCTTCC 3205  
3147 CTTGAGGCTCAAGATCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3206  
3206 TCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3265  
3207 CTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3266  
3266 TGTCTCTCAAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3325  
3267 TGTCTCTCAAAATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3326  
3326 GCTTCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3385  
3327 GCTTCCGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3386  
3386 CTTTCAAGCATTTTGGACTAA 3407  
3387 CTTCAAGACCATCTGGACTGA 3408

Publication No. US20030175967A1

## GENERAL INFORMATION:

APPLICANT: Geron Corporation

APPLICANT: Clark, A. J.

APPLICANT: Denning, Chris

APPLICANT: Cui, Wei

APPLICANT: Zhao, Debbiao

TITLE OF INVENTION: Vectors for Telomerizing Nuclear Donor Cells and Improving the Ef

TITLE OF INVENTION: of Nuclear Transfer

FILE REFERENCE: 732/002

CURRENT APPLICATION NUMBER: US/10/105,616

CURRENT FILING DATE: 2003-01-31

PRIOR APPLICATION NUMBER: US Provisional Application 60/277,749

PRIOR FILING DATE: 2001-03-21

NUMBER OF SEQ ID NOS: 33

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 13766

TYPE: DNA

ORGANISM: Artificial Sequence

## FEATURE:

OTHER INFORMATION: Artificial construct comprising human TERT, myeloproliferative sa

OTHER INFORMATION: rcoma virus (MPSV) promoter, and vector components

-10-105-616-1

Query Match

Best Local Similarity

Matches 2370; Conservative

Score 1593.6; DB 13; Length 13766;

Pred. No. 0;

Mismatches 964; Indels 108; Gaps 9;

35 CACATGACCCCGCTCTCTGTTGCCCCCGGTGGCTCTCTGCTGGCAGCCGATACCG 94

3985 CACCATGCCGCGCTCCCGCTGCTGCGAGCCGCTGGCTCTCTGCTGGCAGCCATACCG 4044

95 GGAGGTGTGGCGCTGGCAACCTTTGTGCGCGCTGGCGCCCGAGGCGCAGGCGGCTGT 154

4045 CGAGGTGTGCTGCTGGCAGCTTCTGCTGGCGCTGGCGCCCGAGGCTGGCGGCTGT 4104

155 GCACCCCGGGACCCGAGATCTACCGACATTTGTTGCTGCTGCTGCTGCTGCTGCTGCT 214

4105 GCAGCGCGGGACCCCGGGGCTTTCCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTG 4164

215 CTGGGCTCAGAGCTCCACCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 274

4165 CTGGGACGACG 4224

275 GCTGTGGCCAGGCTGTGTCAGAGACTCTGCGAGCGCAACGAGAGAAACGCTGTGCTTT 334

4225 GCTGTGGCCAGGCTGTGTCAGAGACTCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 4284

335 TGCTGTGAGCTGCTTAACGAGGCGAGGCGGCGCTCCCATGCGCTTCACTAGTAGCGT 394

4285 CGGCTTCGCGCTGTGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4344

395 GCGTAGCTACTTGGCCCAACACTGTTATTAGAACCTCTGCTGCTGCTGCTGCTGCTGCT 454

4345 GCGCAGCTACTTGGCCCAACACTGTTATTAGAACCTCTGCTGCTGCTGCTGCTGCTGCT 4404

455 ACTGTGAGCGGAGTGGGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514

4405 GCTGTGCGCGCGCTGGGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4464

515 TCTTCTGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574

4465 TGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4524

575 TGCCACCGAGATCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634

4525 CGCTGCCACTCAGGCG 4584

635 CAGGAATTTCATACCTAGTCTTCTTACACAGATCAAGAGCAGTAGTCCGAGGAGC 694

4585 C-----GAACGGGCTGGAAACATAGCTCAGGAGGC 4617

695 ACCGAACCCCTTGGCTTGCATCTCGAGGTACAAAGAGGCATCTGAGTCTCACAGTAC 754  
4618 CGGGTCCCCCTTGGGCTTGCAGCCCGGGTGGAGAGCGGGGGGAGTCCACCGC 4677  
755 AAGTGTGCTTCAAGTAAGAGGCGAGATGCTATCTCTGCTCCCGAGAGTGGAGGAGGACC 814  
4678 AAGTGTGCTTGGTCCCAAGAGGCGCGAGTGGCGCTGCGCTTCTGAGCGGAGCGGAGCC 4737  
815 CACAGGAGAGTGTCTACCAACCCCATCAGGCAAAATCATGGTGTCCCAAGTCTCTGCTCGTC 874  
4738 CGTTGGGCGAGGGGCTCTGGGCGCCACCGGGCGAGAGCGGTGAGCCGAGTACCGCTGTTT 4797  
875 CCCCAGGTGCTACT-----GCAGAGAAGATTTTGTCTTCTTAAAGGAAAGGTGTC 925  
4798 CTGTGTGTGTCTACCTGCGCAGACCGCGCGGAGAGGCCACCTCTTTGGAGGCTGCTCTC 4857  
926 TGACTGTAGTCTCTC---TGGGTGGTGTGCTGTAAACACAAGCCAGCTTCCACAT---C 979  
4858 TGGCAGCGCGCACTCCCAACCATCCGTGGGCGCGCAGCACACCGCGGGCCCCCATCCAC 4917  
980 TCTGCTGTCAACCCCGCCCAAAATGCTTTCAGCTCAGGCCATTTATTGAGACCAAGCA 1039  
4918 ATCGGGCCACACAGCTCCCTGGGACACGCTTGTCCCGGCTGACGCGGAGCAAGCA 4977  
1040 TTTCTTTTACTCCAGGGGAGATGGCCAAAGAGCGTCTAAACCCCTCATTTCTTACTCAGCAA 1099  
4978 CTCTCTTCTCTCAGCGCA---CAAGGAGAGCTGCGGCGCTCTCTTCTTACTCAGCTC 5034  
1100 CTTGAGGCTTACTGCTGGGCGCAGGAGACTGGTGGAGATCATCTTTCTGGGCTCAAG 1159  
5035 TCTGAGGCGGAGCTGCTGGCGCTGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAG 5094  
1160 GCCTAGGACATCAGGACCATCTGTCAGGACACACCGTCTATCCGCTCGATATCTGGCAGAT 1219  
5095 GCTTGTGATGCCAGGACTCCCGCAGGTTGCGCGCTGCGCCAGGCTACTTGGGCAAT 5154  
1220 GCGGCGCTGTTTCAACAGCTGTGTGTAACCATGAGAGTGGCAATATGTACAGCTCT 1279  
5155 GCGGCGCTGTTTCTGAGCTGTGTGGGAACACACGCGCAGTGGCGCTCGGGTGTCTCT 5214  
1280 CAGTCACTTTCAGGTTTCGAACGCAACCAACAGGTGACAGATGCTCT----- 1329  
5215 CAGACGCACTGCGCGCTGCGAGTGGGTCAACCCAGCAGCGGTGTCTGTGCGCGGGA 5274  
1330 -----TGAACACAGCCAGCCACCGCAGCTCAT 1354  
5275 GAAGCCCGAGGCTCTGTGCGCGCGCGGAGGAGGAGACACAGACCCCGCTGCGCTGT 5334  
1355 GATTTGCTGCGCTGACAGAGTCCCTGGCAGGTATATGGTTTCTTTCGGGCTGTCT 1414  
5335 GAGTGTCTGCGCAGCAGCAGCGCTGGCAGCGCTTGGAGGTGTACGGCTTCTGCGGCGCT 5394  
1415 CTGAAGTGTGTGCTGCTAGTCTCTGGGGTACAGGCAATAGCGCGCTTCTTTAA 1474  
5395 GCGCGGCTGTTGCGCGCGCTCTGGGCTCCAGGCAACAGAACGCGCTTCTCTCAG 5454  
1475 GAACTTAAAGTTTCACTCTGTTGGGAAATACGGCAAGCTATCACTCAGGAATGAT 1534  
5455 GAAACCAAGAAGTTTCACTCTCCTGGGAAAGCATGCAAGCTCTCGCTCAGGAGCTGAC 5514  
1535 GTGAAGATGAAGTAGAGGATGCGCTGCGCTCCGACAGCGCGCGGGAAGCGCTGT 1594  
5515 GTGAAGATGAGCGTGGGAGCTGCGCTGCGCTGCGCAGGAGCGCGGAGGTTGGTGTGT 5574  
1595 CCGCTGCGAGACCGCTGAGGAGGAGGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1654  
5575 TCGCGCGCAGACCGCTGCGCTGAGGAGATCTCTGGCAAGTTCTCTGCTGCTGCTGCTGCT 5634  
1655 GGACACATGCTGCTAGAGCTGCTTGGTCACTCTTTTACATCAGAGAGCAGCATTTCA 1714  
5635 GAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5694  
1715 GAAGAACAGGCTCTTCTTCTACCGCTAAGAGTGTGTGGAGCAAGCTGCGAGCATTTGGAGT 1774



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

nucleic - nucleic search, using sw model

1 on: February 1, 2004, 21:42:51; Search time 6853 Seconds  
(without alignments)  
12398.723 Million cell updates/sec

le: US-09-042-460-1  
fect score: 3496  
pence: 1 GAATTCGGGTGGAGGCC.....CCGAGTCGTCACCAAGCTT 3496

ring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

urched: 22781392 seqs, 12152238056 residues

al number of hits satisfying chosen parameters: 45562784

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

EST:

1: em\_estba:\*

2: em\_esthm:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

ult No.	Score	Query Match	Length	ID	Description
1	844.8	24.2	851	13 BU702370	BU702370 UI-M-F10-
2	764.8	21.9	851	12 BG917907	BG917907 602820830
3	655.8	18.8	664	13 BQ258274	BQ258274 NISC_kp11
4	612.4	17.5	641	28 AZ972318	AZ972318 2M0246F07

5	513	14.7	599	10	BB618671	BB618671
6	466	13.3	614	10	BB651920	BB651920
7	440.8	12.6	925	12	BM453198	BM453198 AGENCOURT
8	431.2	12.3	866	10	BE371943	BE371943 601217728
9	347	9.9	347	9	AW244516	AW244516 BR_END068
10	302	8.6	326	13	BY149368	BY149368 BY149368
11	272.2	7.8	492	12	BM824748	BM824748 K-EST0096
12	265	7.6	568	4	BX521269	BX521269 R2PD Mus
13	261.8	7.5	534	9	AW318894	AW318894 un09a02.y
14	235.6	6.7	389	9	AA281296	AA281296 zt08g02.r
15	186.8	5.3	753	13	BU452535	BU452535 603767927
16	172.8	4.9	775	12	BI388013	BI388013 BFL26.002
17	164	4.7	880	13	BU377259	BU377259 603811228
18	160.8	4.6	679	10	BE396606	BE396606 601289077
19	158.8	4.5	610	10	BE314188	BE314188 601316376
20	157.8	4.5	409	9	AA311750	AA311750 EST182469
21	152.2	4.4	715	10	BE396925	BE396925 601290610
22	150	4.3	649	10	BE514070	BE514070 601316575
23	142.6	4.1	696	13	BU139751	BU139751 603134527
24	134	3.8	487	28	AZ799615	AZ799615 2M0057107
25	130.2	3.7	668	14	CA380121	CA380121 659344 NC
26	123.2	3.5	779	10	BE268183	BE268183 601125261
27	118	3.4	875	13	BU122597	BU122597 603148441
28	109.2	3.1	632	14	CA353864	CA353864 625469 NC
29	104.4	3.0	835	13	BU111946	BU111946 603127372
30	101.4	2.9	212	10	BS587267	BS587267 BS587267
31	98	2.8	777	28	AZ826048	AZ826048 2M0101K03
32	97.4	2.8	813	13	BU224024	BU224024 603798349
33	91	2.6	375	9	AA200728	AA200728 mul3h09.r
34	91	2.6	503	9	AI645957	AI645957 mul3h09.y
35	78	2.2	554	28	AQ397020	AQ397020 mgxb0014C
36	76.2	2.2	1230	29	CC211572	CC211572 CH261-14F
37	76.2	2.2	1455	29	CC190951	CC190951 CH261-380
38	76.2	2.2	1554	29	CC190875	CC190875 CH261-38M
39	74.8	2.1	813	10	BG198331	BG198331 RST17599
40	66.8	1.9	739	13	BX315053	BX315053 BX315053
41	66.4	1.9	366	10	BF511837	BF511837 UI-H-B14-
42	64.6	1.8	239	12	BI359552	BI359552 OKST5-7(2
43	63.8	1.8	479	12	BM077893	BM077893 pb21ell.Y
44	63.2	1.8	519	10	BF802688	BF802688 PM4-CI007
45	63.2	1.8	929	29	CNS06T5X	AL414967 T7 end of

## ALIGNMENTS

RESULT 1  
BU702370  
LOCUS  
DEFINITION  
UI-M-F10-byx-f-12-0-UI.r1 NIH-BMAP\_F10 Mus musculus cDNA clone  
IMAGE: 6400523 5', mRNA sequence.  
ACCESSION  
BU702370  
VERSION  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NIH-MGC http://mgc.ncl.nih.gov/  
1 (bases 1 to 851)  
NATIONAL INSTITUTE OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)  
UNPUBLISHED  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

2818	QY	TCGTGGACACTCAGACCTTTGGAGGTTCTCTGTGACTACTCAGGTTATGCCAGACCTCAA	2877
601	Db	TCCTGGACACTCAGACCTTTGGAGGTTCTCTGTGACTACTCAGGTTATGCCAGACCTCAA	660
2878	QY	TTAAGACGAGCCTCACCTTCCACAGTGTCTTCAAAAGCTGGGAAGACCATTGCGGAACAGC	2937
661	Db	TTAAGACGAGCCTCACCTTCCACAGTGTCTTCAAAAGCTGGGAAGACCATTGCGGAACAGC	720
2938	QY	TCCTGTGGGCTTTGCGGTTGAAGTGTCAAGTCTATTTCTAGACTTGCAGGTGACACGCC	2997
721	Db	TCCTGTGGGCTTTGCGGTTGAAGTGTCAAGTCTATTTCTAGACTTGCAGGTGACACGCC	780
2998	QY	TCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTTCAGGCGCTACAGGTTCCATG	3057
791	Db	TCCAGACAGTCTGCATCAATATATACAAGATCTTCTCTGCTTCAGGCGCTACAGGTTCCATG	840
3058	QY	CATGTGTGATT	3068
841	Db	CATGTGTGATT	851

RESULT 2  
 BG917907  
 LOCUS  
 DEFINITION  
 602820830F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:4949887 5',  
 mRNA sequence.

851 bp mRNA linear EST 05-JUN-2001  
 EST 05-JUN-2001

```

SOURCE
  MUS musculus (house mouse)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 851)
  NIH-MGC http://mgc.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Jeffrey Green M.D.
  CDNA Library Preparation: Life Technologies, Inc.
  CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  plate: LLAM10903 row: k column: 08
  High quality sequence stop: 753.
  Location/Qualifiers
    1..851
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="FVB/N"
    /db_xref="taxon:10090"
    /clone="IMAGE:4949887"
    /sex="female virgin"
    /tissue_type="infiltrating ductal carcinoma"
    /dev_stage="5 months"
    /lab_host="DH10B"
    /clone_lib="NCI CGAP Mam6"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sall;
    Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
    Library constructed by Life Technologies. Investigator
    providing samples: Jeffrey Green, M.D., NIH"
  BASE COUNT      184 a 238 c 214 g 215 t
  ORIGIN

Query Match      21.9%; Score 764.8; DB 12; Length 851;
Best Local Similarity 97.5%; Pred. No. 1.2e-197;
Matches 819; Conservative 0; Mismatches 17; Indels 4; Gaps 4;
2582 TGAGGTGACAGCGGATGGTGTGCTTTTACGTTTGTGTGATGACTTCTCTGGTGACGCC 2641

```



```
|||||
9 TGGGTGACGCGGATGGG-TGCTTTTACGTTGG-TGATGACTTTCTGAGGTGAGCC 66
2642 TCACCTGGACCAAGCAAAACCTTCTCAGCACCCCTGGTCCATGGCCCTTCTCAGTATGG 2701
67 TCACCTGGACCAAGCAAAACCTTCTCAGCACCCCTGGTCCATGGCCCTTCTCAGTATGG 126
2702 GTGCATGATAACTTCGACAGACAGTGGTGAATTCCTCTGTGAGCTGGTACCCCTGG 2761
127 GTGCATGATAACTTCGACAGACAGTGGTGAATTCCTCTGTGAGCTGGTACCCCTGG 186
2762 TGGTGCAGCTCCATACCAAGCTGCTCTCAGTGGCTCTTCCCTGTGTGGCTTGTGCT 2821
187 TGGTGCAGCTCCATACCAAGCTGCTCTCAGTGGCTCTTCCCTGTGTGGCTTGTGCT 246
2822 GCACACTCAGCTTTGGAGGTGTTCTGTGACTACTCAGTGTATGCCAGACCTCAATTA 2881
247 GCACACTCAGCTTTGGAGGTGTTCTGTGACTACTCAGTGTATGCCAGACCTCAATTA 306
2882 GACGAGCTCAGCTTTCCAGAGTGTCTTCAAAGCTGGGAAGACCATGGGAACAGCTCT 2941
307 GACGAGCTCAGCTTTCCAGAGTGTCTTCAAAGCTGGGAAGACCATGGGAACAGCTCT 366
2942 GTCCGTCTTGGGTTGAAGTGTACGGTCTATTTCTAGACTTGCAGTGAACAGCTTCCA 3001
367 GTCCGTCTTGGGTTGAAGTGTACGGTCTATTTCTAGACTTGCAGTGAACAGCTTCCA 426
3002 GACGCTCTGCATCAATATATACAGATCTTCTGCTTACAGCTTACAGTGTCCATGCT 3061
427 GACGCTCTGCATCAATATATACAGATCTTCTGCTTACAGCTTACAGTGTCCATGCT 486
3062 TGTGATTCAGCTTCCCTTTGACCGCGTGTAGGAAGAACCTCACAATTTCTGGGCGAT 3121
487 TGTGATTCAGCTTCCCTTTGACCGCGTGTAGGAAGAACCTCACAATTTCTGGGCGAT 546
3122 CATCTCCAGCCAGCATCTGCTGTATGCTATCTGAGTCAAGTCAAGATCCAGGAATGAC 3181
547 CATCTCCAGCCAGCATCTGCTGTATGCTATCTGAGTCAAGTCAAGATCCAGGAATGAC 606
3182 ACTAAGGCTCTGGCTCTCTTCCCTGGAAGCGGCACATTTGGCTCTGTACAGGCTT 3241
607 ACTAAGGCTCTGGCTCTCTTCCCTGGAAGCGG-ACATTTGGCTCTGTACAGGCTT 665
3242 CTT-GCTCAGCTGCTGCTCATTTCTGCTATCTCAAAATGTTCTCTGGGACCTCTGAGGA 3300
666 CTTGCTCAGCTGCTGCTCATTTCTGCTATCTCAAAATGTTCTCTGGGACCTCTGAGGA 725
3301 CAGCCCAAAAACCTGCTGCGGAGCTCCAGAGGCGACAAATGACCATCTTAAAGCTG 3360
726 CAGCCCAAAAACCTGCTGCGGAGCTCCAGAGGCGACAAATGACCATCTTAAAGCTG 785
3361 CAGCTGACCCAGCCTTAAGCAGACTTTTACAGCAATTTTGGACTAACCTGTCTCTTC 3420
786 CAGCTGACCCAGCCTTAAGCAGACTTTTACAGCAATTTTGGACTAACCTGTCTCTTC 845
```

SULT 3  
258274  
FINITION  
BQ258274  
NISC kp11904.q3 Baker mouse embryo e7.5 Mus musculus cDNA clone  
IMAGE:5409222, mRNA sequence.  
BQ258274  
BQ258274.1 GI:20459030  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 664)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
cDNA Library Preparation: J. Baker (Stanford University)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[info@image.llnl.gov](mailto:info@image.llnl.gov)  
MGI:1845958  
Plate: L1AM12043 row: N column: 7  
Seq primer: Sp6 primer.  
Location/Qualifiers  
1...664  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CD-1"  
/db\_xref="taxon:10090"  
/clones="IMAGE:5409222"  
/tissue\_type="embryo, late gastrula"  
/dev\_stage="embryo, 7.5 dpc"  
/lab\_hosts="XLI-Blue"  
/clone\_lib="Baker mouse embryo e7.5"  
/note="vector: pCS105; Site 1: NotI; Site 2: SalI; cDNA  
made by oligo-dT priming. Directionally cloned into  
SalI/NotI sites using the following 5' adaptor:  
5'-TCGACCCAGCGTCCG-3', Size-selected for average insert  
size 1.8-1.9 kb. Library constructed by J. Baker (Stanford  
University)."

FEATURES  
source

BASE COUNT 155 a 177 c 183 g 149 t  
ORIGIN  
Query Match 18.8%; Score 655.8; DB 13; Length 664;  
Best Local Similarity 99.7%; Pred. No. 4.8e-167;  
Matches 557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1103 CCAGCTTAACCTTGAATGGGCGCCAGGAGACTGTGGAGATCATCTTTCTGGGTCAAGGCC 1162  
DB 6 CGAGCTTAACCTTGAATGGGCGCCAGGAGACTGTGGAGATCATCTTTCTGGGTCAAGGCC 65  
QY 1163 TAGACATCAGGACCACTCTGCAGGACACACCGTCTATCGCTCGATCTGGCAGATGCG 1222  
DB 66 TAGACATCAGGACCACTCTGCAGGACACACCGTCTATCGCTCGATCTGGCAGATGCG 125  
QY 1223 GCCCTCTGTTCCAAAGCTGCTGGTGAACCATGACAGAGTGCCAAATATGTCTCAGACTCTCTCAG 1282  
DB 126 GCCCTCTGTTCCAAAGCTGCTGGTGAACCATGACAGAGTGCCAAATATGTCTCAGACTCTCTCAG 185  
QY 1283 GTCACTATGAGGTTTCGAACAGCAACCAAGGTGACAGATGCTTGAACACCAAGGCC 1342  
DB 186 GTCACTATGAGGTTTCGAACAGCAACCAAGGTGACAGATGCTTGAACACCAAGGCC 245  
QY 1343 ACCGCACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 1402  
DB 246 ACCGCACTCATGATTTGCTCCGCTGCACAGCAGTCCCTGGCAGGTATATGTTTCT 305  
QY 1403 TCGGCTCTGCTCTGCAAGTGTGTCTGCTAGTCTCTGGGTACAGGCACAATGAGCG 1462  
DB 306 TCGGCTCTGCTCTGCAAGTGTGTCTGCTAGTCTCTGGGTACAGGCACAATGAGCG 365  
QY 1463 CCGTCTCTTTAAGAACCTTAAGAAAGTTTCATCTGTTGGGGAATACGGCAAGCTATCACT 1522  
DB 366 CCGTCTCTTTAAGAACCTTAAGAAAGTTTCATCTGTTGGGGAATACGGCAAGCTATCACT 425  
QY 1523 GCAGGAACCTGATGGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGCAGGCCCGG 1582  
DB 426 GCAGGAACCTGATGGAAGATGAAGTAGAGATTGCCACTGGCTCCGAGCAGGCCCGG 485  
QY 1583 GAAGGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCTGGTCTACCTTCT 1642  
DB 486 GAAGGACCGTGTCCCGCTGCAGAGCACCGTCTGAGGAGAGGATCTGGTCTACCTTCT 545  
QY 1643 GTTCTGGCTGATGGACACATACGTGTACAGCTGCTTAGGTCAATCTTTTACATCACAGA 1702

|||||  
 546 GTTCTGGCTGATGGACACATACGTGGTACAGCTGTTAGGTCACTTTTACATCACAGA 605  
 1703 GAGCACATTCACAGAGACAGGCTCTTCTTACCGTAAGAGTGTGGAGCAGCTGC 1761  
 606 GAGCACATTCACAGAGACAGGCTCTTCTTACCGTAAGAGTGTGGAGCAGCTGC 664

SULT 4  
 972318  
 TUS  
 FINITION  
 2M0246F07F Mouse 10kb plasmid UUGC2M library Mus musculus genomic  
 clone UUGC2M0246F07 F, genomic survey sequence.

SESSION  
 AZ972318  
 YWORDS  
 AZ972318.1 GI:13843545  
 GSS.

ORCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 641)  
 Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D. Weiss,R.

TITLE  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

JOURNAL  
 MMENT  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA

84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0246 row: F column: 07  
 Seq primer: CGTGTGTAACAGCAGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 641.  
 Location/Qualifiers

1. .641  
 /organism="Mus musculus"  
 /mol type="genomic DNA"  
 /strain="C57BL/6J"  
 /db xref="taxon:10090"  
 /clone="UUGC2M0246F07"  
 /sex="female"  
 /lab host="E. coli strain XL10-Gold, Tl-resistant, F-"  
 /clone lib="Mouse 10kb plasmid UUGC2M library"  
 /notes="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (female) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (GI|4732114|gb|AP129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

155 a 195 c 157 g 134 t

SE COUNT  
 IGIN

Query Match 17.5% Score 612.4; DB 28; Length 641;  
 Best Local Similarity 99.0% Pred. No. 3.1e-155;  
 Matches 616; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 762 CTTTCAGCTAAGAGGCGCAGATGCTATCTCTGTCGCCAGAGTGGAGGAGGCCCCACAGG 821  
 DB 20 CATGAATCGCAGAGGCGCAGATGCTATCTCTGTCGCCAGAGTGGAGGAGGCCCCACAGG 79

QY 822 CAGTGTGTACCAACCCCATCAGGCAATCATGGTGCACAGTCCCTGCTCGTCCCGGAG 881  
 DB 80 CAGTGTGTACCAACCCCATCAGGCAATCATGGTGCACAGTCCCTGCTCGTCCCGGAG 139

QY 882 GTGCTACTGTCAGAGGAGAGTGTCTTCTTAAAGGAAGGTGTCTGACCTGAGTCTCTCT 941  
 DB 140 GTGCTACTGTCAGAGGAGAGTGTCTTCTTAAAGGAAGGTGTCTGACCTGAGTCTCTCT 199

QY 942 GGGTCGGTGTCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACACCCGCCAA 1001  
 DB 200 GGGTCGGTGTCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACACCCGCCAA 259

QY 1002 AATGCTTTGAGCTCAGGCCATTTATTGAGACCAGACATTTCTTTTACTCCAGGGAGAT 1061  
 DB 260 AATGCTTTGAGCTCAGGCCATTTATTGAGACCAGACATTTCTTTTACTCCAGGGAGAT 319

QY 1062 GGCACAAGAGCGCTTAAACCCCTCATTTCTACTCAGCAACCTCCAGCCTTAACCTGCG 1121  
 DB 320 GGCACAAGAGCGCTTAAACCCCTCATTTCTACTCAGCAACCTCCAGCCTTAACCTGCG 379

QY 1122 GCCAGGAGCTGTGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCCTC 1181  
 DB 380 GCCAGGAGCTGTGTGGAGATCATCTTTCTGGCTCAAGGCTAGGACATCAGGACCCTC 439

QY 1182 TGCAGGACACACCGCTCTATCGCTCGATATCTGGCAGATGCGGCCCTGTTCACACAGCTG 1241  
 DB 440 TGCAGGACACACCGCTCTATCGCTCGATATCTGGCAGATGCGGCCCTGTTCACACAGCTG 499

QY 1242 CTGTGTGAACCATCAGAGTGCCTAATATGTGAGACTCTCTCAGGTCAATTCAGGTTTGA 1301  
 DB 500 CTGTGTGAACCATCAGAGTGCCTAATATGTGAGACTCTCTCAGGTCAATTCAGGTTTGA 559

QY 1302 ACAGCAAAACCAACAGGTGACAGATGCTTGAACACAGCCACCGACCTCATGGATTG 1361  
 DB 560 ACAGCAAAACCAACAGGTGACAGATGCTTGAACACAGCCACCGACCTCATGGATTG 619

QY 1362 CTCCTGCTGACAGAGTCCCT 1383  
 DB 620 CTCCTGCTGACAGAGTCCCT 641

RESULT 5  
 BB618671  
 LOCUS  
 BB618671 RIKEN full-length enriched, 8 days embryo Mus musculus  
 cDNA clone 5730412M20 5', mRNA sequence.

ACCESSION  
 BB618671  
 VERSION  
 BB618671.1 GI:16458173  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)

ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 599)

REFERENCE  
 AUTHORS  
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,  
 Hiranoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda  
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,J., Ohno,M., Sasaki  
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki  
 D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,  
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,  
 Muramatsu,M. and Hayashizaki,Y.

TITLE  
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)  
 JOURNAL  
 Unpublished  
 COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/  
M., Konno, H., Okazaki, Y., Muramatsu, M., Sugahara, Y., Shibata, K., Itoh  
Carninci, P., Hayashizaki, Y., Muramatsu, M., and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
Y., and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamana, I., Aizawa  
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
Hayashizaki, Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

# FEATURES

Location/Qualifiers  
1. 599  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="5730412M20"  
/sex="mixed"  
/dev\_stage="8 days embryo"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 8 days embryo"  
/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGGAGAGGACGACGAGCTCTTTTCTTTTCTT 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of subtraction to  
Ret = 100.0 Second strand cDNA was prepared with the  
primer adapter of sequence [5'  
GAGAGGAGATCTCGATTCTAATAATATCCCGCCCCCCCC 3']. cDNA  
was cloned into the XhoI and BamHI sites. Vector: a  
modified pBluescript KS(+) after bulk excision from Lambda  
FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."  
SE COUNT 91 a 185 c 181 g 142 t  
IGIN  
Query Match 14.7%; Score 513; DB 10; Length 599;  
Best Local Similarity 100.0%; Pred. No. 3.5e-128; Indels 0; Gaps 0;  
Matches 513; Conservative 0; Mismatches 0;  
10 GTGGAGGCGCCATCCCGCTTGTAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGC 69  
86 GTGGAGGCGCCATCCCGCTTGTAGCACAATGACCGCGCTCTCTGTTGCCCGCGGTGC 145  
70 GCTCTCTGTCGCCACCGATACCGGAGGTGGCGCTGGCAACCTTTTGGCGGCC 129

Db 146 GCTCTCTGTCGCCACCGATACCGGAGGTGTGGCGCTGGCAACCTTTTGGCGGCC 205  
QY 130 TGGGCGCCGAGGCGAGGGCGGTGTGTGCAACCGGAGACCGGAGATCTACGCACTTTGG 189  
Db 206 TGGGCGCCGAGGCGAGGGCGGTGTGTGCAACCGGAGACCGGAGATCTACGCACTTTGG 265  
QY 190 TTGCCCAATGCTCTAGTGTGCATGCACTGGGGCTCACAGCTCCACCTGCCGACCTTTCT 249  
Db 266 TTGCCCAATGCTCTAGTGTGCATGCACTGGGGCTCACAGCTCCACCTGCCGACCTTTCT 325  
QY 250 TCCACAGGTGTCTATCCCTGAAAGAGCTGTGTGGCCAGGGTTGTGCAGAGACTCTGCGAGC 309  
Db 326 TCCACAGGTGTCTATCCCTGAAAGAGCTGTGTGGCCAGGGTTGTGCAGAGACTCTGCGAGC 385  
QY 310 GCAACGAGAGAAACGCTGTGCTTTTGGCTTTGAGCTGTAAACGAGGCCAGAGGGCGGCG 369  
Db 386 GCAACGAGAGAAACGCTGTGCTTTTGGCTTTGAGCTGTAAACGAGGCCAGAGGGCGGCG 445  
QY 370 TCCCATGGCTTCACTAGTAGCTGCTACTTGTCCCAACACTGTTATTGAGACCC 429  
Db 446 TCCCATGGCTTCACTAGTAGCTGCTACTTGTCCCAACACTGTTATTGAGACCC 505  
QY 430 TCGCTGTCACTGTGTGCATGCTACTTGTGAGCGAGTGGCGGAGCTGCTGTCT 489  
Db 506 TCGCTGTCACTGTGTGCATGCTACTTGTGAGCGAGTGGCGGAGCTGCTGTCT 565  
QY 490 ACTCTGTGGCACACTGTGCTCTTTTATCTTTCTGG 522  
Db 566 ACTCTGTGGCACACTGTGCTCTTTTATCTTTCTGG 598

## RESULT 6

BB651920 614 bp mRNA linear EST 26-OCT-2001  
BB651920 RIKEN full-length enriched, ES cells Mus musculus cDNA  
clone C30020G14 5', mRNA sequence.  
BB651920  
BB651920  
BB651920.1 GI:16486059  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 614)  
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
Hiramoto, K., Hori, F., Iehli, Y., Ito, M., Kawai, J., Konno, H., Kouda  
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki  
D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
Unpublished  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsc.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)  
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)  
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawaji, J., Shibata, K. and Hayashizaki, Y.  
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)  
Please visit our web site (<http://genome.scri.riken.go.jp>) for further details.

```
. mouse tissues.  
Location/Qualifiers  
. i .614
```

/organism="Mus musculus"  
/mol\_type="mRNA "  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="C33002OGI4"  
/cell\_type="ES cells"  
/lab\_host="SOLR"  
/clone\_lib="RIKEN full-length enriched, ES cells"  
/note="Site 1: XhoI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in RIKEN contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']  
  
GAGAGAGAGAAGATCCCAAGAGCTTTTTTTTATTTCVN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand GABA was prepared with the primer adapter of sequence [5'  
cGAGAGAGAAATCTCGAGTTAATAATTAATCCCCCCC C 3' ]."

93 a    191 c    188 g    142 t

BASE COUNT  
-----  
ggttgc

Query Match.	13.3%;	Score 466;	DB 10;	Length 614;
Best Local Similarity	99.8%;	Pred. No. 2.3e-115;		
Matches 477;	Conservative 0;	Mismatches 1;	Indels 1;	Gaps 1;
10	GTGGGAGGCCATCCCGGCTTGAGCACAATGACCCGCGCTCTCGTTGGTCCCGCGGTGC	69		
86	GTGGGAGGCCATCCCGGCTTGAGCACAATGACCCGCGCTCTCTGTTCCCGCGGTGC	145		
70	GCTCTCTGTGTGCGACGCCBATAACGGGAGGTGTGGCGCTGGCAACCTTTGTGGGGGCC	129		
146	GCCTCTGTGTGCGACGCCGATACCGGAGGTGTGGCGCTGGCAACCTTTGTGGGGGCC	205		
130	TGGGGCCGAGGGCAGGGCGCTTGCAACCCGGGACCCGAGATCTACCGCACTTTGG	199		
206	TGGGGCCGAGGGCAGGGCGCTTGCAACCCGGGACCCGAGATCTACCGCACTTTGG	265		
190	TTGCCCAATGCCTAGTGTGCACTGACCTGGGGCTCACAGCCTCCACCTGCCGACCTTTCT	249		
266	TTGCCCAATGCCTAGTGTGCACTGACCTGGGGCTCACAGCCTCCACCTGCCGACCTTTCT	325		
250	TCACCAAGTGTCAATCCCTGAAGAGCTGTGTGGCAGGGTGTGCAAGATCTTCGAGC	309		
326	TCACCAAGTGTCAATCCCTGAAGAGCTGTGTGGCAGGGTGTGCAAGATCTTCGAGC	385		
310	GCAACGAGAGAAACGTGCTGGGCTTTTGGCTTTGAGCTGCTTAACGAGGSCCAGAGCGGGC	369		
386	GCAACGAGAGAAACGTGCTGGGCTTTTGGCTTTGAGCTGCTTAACGAGGSCCAGAGCGGGC	445		
370	CTCCCATGGCTTCACTAGTAGCGTGCAGTACTCTTGGCCCAACACTGTTATGAGACC	429		
446	CTCCCATGGCTTCACTAGTAGCGTGCAGTACTCTTGGCCCAACACTGTTATGAGACC	505		

[illegible]

FEATURES	source	Location/Qualifiers
1.	925	/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IWAGR:5529940"
		/tissue_type="leiomyosarcoma"
		/lab_host="DH10B (phage-resistant)"
		/clone_lib="NIH MGC 71"
		/notes="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT
		Average insert size 2.1 kb. "
BASE COUNT	172 a	290 c 276 g 187 t
ORIGIN		
	Query Match	12.6%; Score 440.8; DB 12; Length 925;
	Best Local Similarity	73.0%; Pred. No. 2.1e-108;
	Matches	597; Conservative 0; Mismatches 212; Indels 9; Gaps 2;
QY	2103	GCTGTGACACAGACACCCAGGATGTACTTTGTATTAGGCAGATGTACCGGGGCGCTATGAT 2162
Db	3	GCCAGGACCGCGCGCTGACGTGTACTTTGTCAAGGTGGATGTACGGGGCGGTAGCAC 62
QY	2163	GCCATCCCCCAGGTAAGCTGGTGGAGGTGTGGCAATATGATCAGGCACTCGGAGAGC 2222
Db	63	ACCATTTCCAGGACAGGCTCACGGAGGTATCGGCAGATCATCA---AACCCAGAAC 119
QY	2223	ACGTACTGTATCCGCCAGTATGCAGTGTGTCGGAGAGATAGCAAGGCCAAAGTCCACAAG 2282
Db	120	ACGTACTCGTGGTCCGTATGCGTGTGTCAGAGGCCGCCCATGGCACGTCCGCAAG 179
QY	2283	TCTTTTAGGAGACAGGTCACACACCTCTCTGACCTCCAGCCATACATGGGCGAGTTCCTT 2342
Db	180	GCTTTCAAGACCAAGTCTCTACCTTTGACAGACCTCCAGCCGTATACATGCGAGTTCGTG 239
QY	2343	AAGCATCTGCAGGATTCCAGATGCCAGTGCACCTGAGGAATCCGGTTGTTCATCGAGCAGAGC 2402
Db	240	GCTACCTGCGAG-----GAGACACGCCGCTGAGGGATGCCGTGTCATTCAGCGCAGAGC 293
QY	2403	ATCTCTATGATGAGACAGCAGCAGCTGTTTGAATCTTCTCTGCATCTCTCGGTCCAC 2462



	Matches	347;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2227	ACTGTATCCCGCAGTATG	CAGTGGTCCGAGAGATAG	CCAGGCCAAGTCC	CAAGTCC	2286				
Db	347	ACTGTATCCCGCAGTATG	CAGTGGTCCGAGAGATAG	CCAGGCCAAGTCC	CAAGTCC	288				
Qy	2287	TTAGGAGACAGGTG	CACCAACCTCTCTG	ACCTCCAGCCATAT	CATGGGCCAGTCTT	2346				
Db	287	TTAGGAGACAGGTG	CACCAACCTCTCTG	ACCTCCAGCCATAT	CATGGGCCAGTCTT	228				

2347	ATCTGCAGGATTCAGATGCCAGTGCACCTGAGAACTCCGTTGTTCATCGAGCAGAGCATCT	2400
Qy		
227	ATCTGCAGGATTCAGATGCCAGTGCACCTGAGAACTCCGTTGTTCATCGAGCAGAGCATCT	168
Db		
2407	CTATCAATGAGAGCAGCAGCAGCCTGTTTGACTTCTCTGCACCTTCTCGGTGCACAGTG	2466
Qy		
167	CTATCAATGAGAGCAGCAGCAGCCTGTTTGACTTCTCTGCACCTTCTCGGTGCACAGTG	108
Db		
2467	TCGTAAGAATTGGTGAACAGTGCTATACGACGTGCCAGGGCATCCCCAGGGCTCCAGCC	2526
Qy		
107	TCGTAAGAATTGGTGAACAGTGCTATACGACGTGCCAGGGCATCCCCAGGGCTCCAGCC	48
Db		
2527	TATCCACCCCTGCTCGACGTCTGTGTTTCGGAGACATGGAGAAACAAG	2573
Qy		
47	TATCCACCCCTGCTCGACGTCTGTGTTTCGGAGACATGGAGAAACAAG	1
Db		

RESULT 10	BY149368	326 bp	linear	EST 10-DEC-2002					
LOCUS	BY149368	RIKEN full-length enriched, 17.5 days embryo whole body							
DEFINITION	Mus musculus cDNA clone L930261L24 5', mRNA sequence.								
ACCESSION	BY149368								
VERSION	BY149368.1	GI:26285897							
KEYWORDS	EST.								
SOURCE	Mus musculus (house mouse)								
ORGANISM	Mus musculus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
AUTHORS	1 (bases 1 to 326) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,								

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojbori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Cruckenbush, J., Schriem, L. M., Kanpin, A., Matsuda, H., Batalov, S., Betsel, K. W., Blake, J. A., Bratt, D., Brusic, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Malcais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perleac, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, K., Verardo, R., Wagner, J., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yangisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hata, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, S. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

JOURNAL

TITLE

LABORATORY PRESS, COCA SPRING HARBOR, NEW YORK.

78 a	88 c	100 q	81 b
------	------	-------	------

12466831  
Contact: Yoshihide Hayashizaki

**Contact:** Yoshinori Hayashizaki  
**Laboratory for Genome Exploration Research Group, RIKEN Genomic**

Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp/  
URL: http://genome-gsc.riken.go.jp/  
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane  
'T., Imotani, K., Itoh, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A.,  
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,  
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,  
'M., Waki, K., Watanabe, A., Watanabe, M., and Hayashizaki, Y. Direct  
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN Integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multichannel sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in Riken  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Please visit our web site (http://genome-gsc.riken.go.jp) for  
further details.

## FEATURES

## Source

Location/Qualifiers

1..326  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6J"  
/db\_xref="taxon:10090"  
/clone="U930261124"  
/class="whole body"  
/dev\_stage="17.5 days embryo"  
/clone\_1db="RIKEN full-length enriched, 17.5 days embryo  
whole body"

## BASE COUNT

58 a 92 c 98 g 78 t

## ORIGIN

## Query Match

Best Local Similarity 98.1%; Score 302; DB 13; Length 326;  
Pred. No. 7.2e-71; Mismatches 5; Indels 1; Gaps 1;  
Matches 316; Conservative 0;

236 TGCCGACCTTCTCCACAGGAGTGTATCCCTGAAAGAGCTGTGGCCAGGTTGTGCA 295  
1 TGCCGACCTTCTCTCCACAGGAGTGTATCCCTGAAAGAGCTGTGGCCAGGTTGTGCA 60  
296 GAGACTCTGCGAGCGCAACGAGAGAAAGTGTGCTTTTGGCTTTGAGCTGTTCACGA 355  
61 GAGACTCTGCGAGCGCAACGAGAGAAAGTGTGCTTTTGGCTTTGAGCTGTTCACGA 120  
356 GAGCGAGGCGGCGCTCCATGCTTCACTAGTAGGCGTGTAGTACTCTG-CCCAAA 414  
121 GAGCGAGGCGGCGCTCCATGCTTCACTAGTAGGCGTGTAGTACTCTG-CCCAAA 180  
415 CTGTTATTTAGACCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 474  
181 CTGTTATTTAGACCTGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 240  
475 AAGACCTGCTGTCTACTCTGTGGCAACGTCGTTTATCTTTTGGGCCCCCGAGCT 534  
241 AAGACCTGCTGTCTACTCTGTGGCAACGTCGTTTATCTTTTGGGCCCCCGAGCT 300  
535 GTGCTTACAGAGTGTGTGGTCT 556  
301 GTGCTTACAGAGGAGATGGCC 322

## RESULT 11

BM824748 492 bp mRNA linear EST 06-MAR-2002  
LOCUS BM824748  
DEFINITION K-EST0096335 S22SNUI6n1 Homo sapiens cDNA clone S22SNUI6n1-99-E07  
5', mRNA sequence.

## ACCESSION

BM824748  
BM824748.1 GI:19181161  
EST.

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 492)

## AUTHORS

Oh, K. J., Cheong, J. E., Sohn, H. Y., Kim, J. M., Park, H. S., Kim, S. and  
Kim, Y. S.

## TITLE

21C Frontier Korean EST Project 2001  
Unpublished  
Contact: Kim YS

## JOURNAL

Genome Research Center

## COMMENT

Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 99 row: E column: 07  
High quality sequence stop: 492.

## FEATURES

## Source

Location/Qualifiers

1..492  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S22SNUI6n1-99-E07"  
/sex="F"  
/tissue\_type="Asclites"  
/cell\_type="Lymphoblast-like"  
/cell\_line="SNU-16"  
/lab\_host="DH10B"  
/lab\_host="S22SNUI6n1"  
/note="Organ: Stomach; Vector: pT73-Pac; Site: 1; EcoRI;  
Site 2: NotI; The S22SNUI6 library was contributed by the  
Seared laboratory and it was constructed as described by  
Bonald, M. F., Lennon, G. and Soares, M. B. (1996), Genome  
Research 6(9): 791-806. RNA was prepared from harvested  
cells of SNU-16 culture. SNU-16 cell was obtained from  
Korean Cell Line Bank (KCLB). SNU-16 was established from  
ascitic fluids of Korean patients by Park J. G. et al.  
(1990), Cancer Res 50: 2773-2780."

## BASE COUNT

96 a 152 c 131 g 113 t

## ORIGIN

## Query Match

Best Local Similarity 75.7%; Score 272.2; DB 12; Length 492;  
Pred. No. 1.2e-62; Mismatches 108; Indels 0; Gaps 0;  
Matches 337; Conservative 0;

2730 GTGAATCTCCCTGTGAGCTGTATCCCTGAGTGTGAGCTCATACAGCTGCTCT 2789  
1 GTGAATCTCCCTGTGAGCTGTATCCCTGAGTGTGAGCTCATACAGCTGCTCT 60  
2790 CACTGCTGTTTCCCTGTGAGCTGTATCCCTGAGTGTGAGCTCATACAGCTGCTCT 2849  
61 CAGGCGCTATTTCCCTGTGAGCTGTATCCCTGAGTGTGAGCTCATACAGCTGAGAGC 120  
2850 GACTACTCAGGTATGAGCGAGACCTCAATTAAGAGAGCTCACCTTCAGAGTCTTC 2909  
121 GACTACTCAGGTATGAGCGAGACCTCAATTAAGAGAGCTCACCTTCAGAGTCTTC 180  
2910 AAGCTGGAGAGCAATGCGAGAAAGAGCTCTGTGCTTGGGTTGAAGTGTACGGT 2969  
181 AAGCTGGAGAGCAATGCGAGAAAGAGCTCTGTGCTTGGGTTGAAGTGTACGGT 240  
2970 CTATTCTAGACTTGAGAGTGAAGAGCTCCAGAGAGTGTGATGAATATATACAGATC 3029







## REFERENCE

1 (bases 1 to 753)

Boardman, P.E., Samz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,  
Fong, W.T., Tickie, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

## TITLE

A Comprehensive Collection of Chicken cDNAs

## JOURNAL

Curr. Biol. 12 (22), 1965-1969 (2002)

## MEDLINE

2235534

12445392

Contact: Simon Hubbard

## COMMENT

Department of Biomedical Sciences  
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QP, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

## FEATURES

## Source

1..753

/organism="Gallus gallus"

/mol\_type="mRNA"

/strain="Layer"

/db\_xref="taxon:9031"

/clone="CHST696f7"

/sex="Female"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="CSEORNI14"

/note="Organ: ovary; Vector: Bluescript II KS(+); Site\_1:  
EcoRI; Site\_2: NotI; This normalized library was  
constructed from 1 million independent clones. cDNA  
synthesis was initiated using an oligo(dT) primer, using  
methylated C in the first strand synthesis reaction.  
Following this first strand reaction, double-stranded cDNA  
was blunt-ended, ligated to NotI adapters, digested with EcoRI  
, size-selected, and cloned into the NotI and EcoRI  
compatible sites of a custom modified MCS of the  
pBluescript (KS+) vector. The library was normalized in 2  
rounds using conditions adapted from Soares et al., PNAS  
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6  
(1996): 791, except that a significantly longer  
reannealing hybridization was used."

## BASE COUNT

230 a 155 c 132 g 236 t

## ORIGIN

## Query Match

Best Local Similarity 5.3%; Score 186.8; DB 13; Length 753;

## Matches 312; Conservative 0; Mismatches 177; Indels 9; Gaps 1;

```

QY 2919 AAGACCATGCGAACAAGCTCTGCGGCTTGAAGTGCAGGTCATTTCTA 2978
DB 5 AAAAATGAAATGCAATTAAGTCAAGTCCCAAGTGAATGCCATCCTTACTT 64
QY 2979 GACTTGAGGTGAACAGCTCCAGACAGTCTGCATCATATATACAGATCTTCTGCTT 3038
DB 65 GACTTAAAGATCAACAGCTTCAAGACAGTCTTAATTAACATCAAGATATTTTACTT 124
QY 3039 CAGGCTACAGGTTCATGATGATTCAGTCCCTTTCAGACAGGTTAGGAAG 3098
DB 125 CAGGCTTACAGTTCATGCTGTGTCTTCAAGCTTCATCAACAGAAAGTTAGGAAT 184
QY 3099 AACCTCACATCTTTTGGGACATCTCCAGCAAGCATCTGCTGTAATGCTATCTG 3158
DB 185 AATCTGATTTCTTCTTAAGATCATCTGATGATGCTTCATGCTGCTATTTATCTG 244
QY 3159 AAGTCAAGATCCAGAAATGACACT-----AAAGGCTCTGAGCTCTTCTCTCT 3209
DB 245 AAAGCTAAATCAAGAGTTCTTTAGTAGCAAGATGATCTGTATGTTCCCTTTT 304
QY 3210 GAAGCCGACATGCTCTGCTACAGAGCTTCTGCTCAAGCTGCTGCTCATCTGTC 3269
DB 305 GAGGACAGCAAAATGGCTGTGCTACATGCTTCATGTCAACTGCCAACACAAAGTT 364
QY 3270 ATGTACAAATGTCCTGAGGACCTGAGGACAGCCAAAACCTGCTGTGCGGAAGCTC 3329

```

```

DB 365 ATTACAAATGCTTACTTAGCCCCCTTAAGTCTAATAGATGATCTGTGGAAGATC 424
QY 3330 CCAGAGCGACATGACCATCTTAAAGCTGAGCTGACCCGACCTTAGACACAGACTTT 3389
DB 425 CCAAGGATCTATGAACTGCTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 484
QY 3390 CAGACATTTGAGACTTA 3407
DB 485 AAATATAGGACTTA 502

```

Search completed: February 2, 2004, 03:44:55  
Job time : 6855 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 1, 2004, 15:05:59 ; Search time 9856 Seconds

(without alignments)  
4657.123 Million cell updates/sec

Title: US-09-042-460-2

Perfect score: 5901  
Sequence: 1 MTRAPRCPRVRSLSRSRYRE.....TLKAAADPALSTDFQTILD 1122

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n model -DEV=xlh  
-Q=cgn2.1/usfpo.spool/us09042460/runat.01022004.150551.17014/app.query.fasta\_1.1287  
-DB=GenEmbl -QFMT=fastap -SUPFIX=rgs -WINDMATCH=0.1 -LOOPEXT=0 -LIST=45  
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -ICPU=3  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=45 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USRF=US09042460@cgn2.1.6581@runat.01022004.150551.17014 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFRY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl:.\*  
1: gb\_ba:.\*  
2: gb\_hgt:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_str:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vi:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_str:.\*  
28: em\_un:.\*

29: em\_vi:.\*  
30: em\_hgt\_hum:.\*  
31: em\_hgt\_inv:.\*  
32: em\_hgt\_other:.\*  
33: em\_hgt\_mus:.\*  
34: em\_hgt\_pun:.\*  
35: em\_hgt\_rtd:.\*  
36: em\_hgt\_mam:.\*  
37: em\_hgt\_vrt:.\*  
38: em\_sy:.\*  
39: em\_hgt\_hum:.\*  
40: em\_hgt\_mus:.\*  
41: em\_hgtg\_other:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5901	100.0	3369	10 AF073311	AF073311 Mus muscu
2	5901	100.0	3426	10 AF051911	AF051911 Mus muscu
3	4751	80.5	4170	10 AF149012	AF149012 Mesocric
4	3505	59.4	3396	6 BD091553	BD091553 Adult bon
5	3505	59.4	3396	6 BD094749	BD094749 The cell
6	3505	59.4	3396	6 BD096291	BD096291 Cells cap
7	3505	59.4	3399	6 AX481414	AX481414 Sequence
8	3505	59.4	3399	6 AX481414	AX481414 Sequence
9	3505	59.4	4015	6 AR104587	AR104587 Sequence
10	3505	59.4	4015	6 AR175848	AR175848 Sequence
11	3505	59.4	4015	6 AR182221	AR182221 Sequence
12	3505	59.4	4015	6 AR224455	AR224455 Sequence
13	3505	59.4	4015	6 AR226390	AR226390 Sequence
14	3505	59.4	4015	6 AR243328	AR243328 Sequence
15	3505	59.4	4015	6 AR263555	AR263555 Sequence
16	3505	59.4	4015	6 AR265996	AR265996 Sequence
17	3505	59.4	4015	6 AX019310	AX019310 Sequence
18	3505	59.4	4015	6 AX019310	AX019310 Sequence
19	3505	59.4	4015	6 AX526695	AX526695 Sequence
20	3505	59.4	4015	6 BD011044	BD011044 Human tel
21	3505	59.4	4015	6 BD011832	BD011832 Quantitat
22	3505	59.4	4015	6 BD082985	BD082985 Method fo
23	3505	59.4	4015	6 BD131727	BD131727 Method fo
24	3505	59.4	4015	6 E36793	E36793 Human telom
25	3505	59.4	4015	6 AF015950	AF015950 Homo sapi
26	3505	59.4	4015	6 AX001446	AX001446 Sequence
27	3505	59.4	4015	6 AX003121	AX003121 Sequence
28	3505	59.4	4015	6 BD136185	BD136185 Human tel
29	3505	59.4	4015	6 AX391846	AX391846 Sequence
30	3505	59.4	4015	6 AF018167	AF018167 Homo sapi
31	3505	59.4	4015	6 AF043739	AF043739 Synthetic
32	3505	59.4	4015	6 BD011070	BD011070 Human tel
33	3505	59.4	4015	6 E36819	E36819 Human telom
34	3505	59.4	4015	6 AX553919	AX553919 Sequence
35	3505	59.4	4015	6 AX553924	AX553924 Sequence
36	3505	59.4	4015	6 AR104586	AR104586 Sequence
37	3505	59.4	4015	6 AR175847	AR175847 Sequence
38	3505	59.4	4015	6 BD011069	BD011069 Human tel
39	3505	59.4	4015	6 E36818	E36818 Human telom
40	3505	59.4	4015	6 AR243330	AR243330 Sequence
41	3505	59.4	4015	6 BD011046	BD011046 Human tel
42	3505	59.4	4015	6 E36795	E36795 Human telom
43	3505	59.4	4015	6 AB085628	AB085628 Homo sapi
44	3505	59.4	4015	6 AF247818	AF247818 Rattus no
45	3505	59.4	4015	6 AB086379	AB086379 Homo sapi
	2919	49.5	2992	9 AB086590	AB086590 Homo sapi

#### ALIGNMENTS

RESULT 1

AF073311 3369 bp mRNA linear ROD 09-SRP-1998  
 LOCUS AF073311  
 DEFINITION Mus musculus telomerase catalytic subunit mRNA, complete cds.  
 ACCESSION AF073311  
 VERSION AF073311.1 GI:3551846  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 3369)  
 Martin-Rivera, L., Herrera, E., Albar, J. P. and Blasco, M. A.  
 Expression of mouse telomerase catalytic subunit in embryos and  
 adult tissues  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (18), 10471-10476 (1998)  
 MEDLINE 98393668  
 PUBMED 9724727  
 REFERENCE  
 2 (bases 1 to 3369)  
 Martin-Rivera, L., Herrera, E. and Blasco, M. A.  
 Direct Submissions  
 Submitted (19-JUN-1998) Immunology and Oncology, National Centre of  
 Biotechnology, Cantoblanco, Madrid 28049, Spain  
 FEATURES  
 source location/qualifiers  
 1..3369  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1..3369  
 /codon\_start=1  
 /product="telomerase catalytic subunit"  
 /protein\_id="AAC34821.1"  
 /db\_xref="GI:3551847"  
 /translation="MTRAPRCPAVRSLRSRYEWMPLATVRLQPEGRRLVQDGP  
 KIRYLVAQCIVCMGSGOPADLSFHSQVSLKEIVARVQRLCERNRNLVLAQFE  
 ILNEARGPMAFTSVSVSYLPNTVETLVSGAMMLLSRVDDLLVTLVLAHAGFE  
 LVPSGACVCGSPLOYICATTDIMPSSVSYRPTRGVGNFNLAFLQIKSSSGOE  
 APKPLAPRGCTKRLSLTSTSVSAKACVCPUPREBGNQVLPISGASWSPDS  
 ARSEVPFAEKDLSKGVSDLSGVCCKMRPSSTLSLSPKQNAFQRLPFIETRRH  
 FLYSRGDQERLNPSTLSTLQPNLTARRLVETIFGSSPRTSGLPCTHRLSRVW  
 QMRPLFQQLLVNAHCQVRLRLSHRCFRANQVDALNTSPHMLDLRLSHWQ  
 VYGFRAICLVVASIMGTRHNRFPFKLKFISLIGKYGKLSLOELMMKXVEDCH  
 WLRSRPGDQRPVPAEHLREIRLILATFLFMDTYVOLLSFYIESTFQKRLFFY  
 KSVWSKQSIQSVROHLERVLRLSGEBYRHODTWLAMPICRLAFIPKNGRLIV  
 NMSVSMGTRALGRKQKQHPQRLKTLPSMLYERTHPHLMGSSVTKGNDLYRTMRA  
 FVLRALDQTPRMTEVKAQVDTGYADLPQKLYEVVYAMTRISESTCTROYAVRR  
 DSQGVHRSFRQVTTLSDDQPIWQFLKLDQSDSLNSVLIQSISMWSSSL  
 FDFLFLRHSVVKIGRCYQCGILQSGSLSTLCSLFGMEKMLFAVQRGSL  
 LRFVDDFLVTPHLDQATKFLSTLVHGVPEYCGIMNLQKTVNFPEVPTLGGAAVQ  
 LPAHCLPWCGLLDLTQLEVCDSVGAQTSIKTSLFQSVKAKTKENKLSYLR  
 LKCHGLFLDLOVNSLQTVGINIKYIFLLOVAPACVLCIPFORVRKMLTFPIGIT  
 SOASCCVAILIKVKNMGKTKASGSPPEAAHWTCYQAFLKLAHSAVITYKLLGLPRT  
 AOKLRCRLPFAVNTILKAAADPLSDFTQTLID"

## CDS

BASE COUNT 737 a 959 c 891 g 782 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.63e-299 Length: 3369  
 Score: 5901.00 Matches: 1122  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-042-460-2 (1-1122) x AF073311 (1-3369)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 Db 1 ATGACCCCGGCTCCGCTTGCCGCCGCGTGGCTCTCTGCGCCAGCCATCCCGAGG 60  
 QY 21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgValGln 40  
 Db 61 GTGTGGCCGCTGGCACTTTGTGGGGCGCTGGGGCCGAGGCGAGCGCGCTTGTCGA 120

QY 41 ProGlyAspProIleIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
 Db 121 CCGGGGAGCCCGAAGATCTACCGCACTTGGTTGCCCAATGCTGATGATGACACG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHISGlnValSerSerLeuYsgLeu 80  
 Db 181 GGCTCAGCCCTCAGCTCCGACCTTCTTCACACAGGTCTCATCCCTGAAAGAGCTG 240  
 QY 81 ValAlaArgValAlaGlnArgLeuGlyGluValArgGlnValArgAspValLeuAlaPheGly 100  
 Db 241 GTGGCAGGAGTGTGAGAGACTTCCGAGCCGACAGAGAAAGTGTGCTTGTGGC 300  
 QY 101 PheGlnLeuLeuGlnAlaArgGlyGlyProPrometAlaPheThrSerSerValArg 120  
 Db 301 TTGAGGCTGTTAAGAGGACAGAGGCGGGCTCCCATGCTTCATAGTACGCTGCT 360  
 QY 121 SerTyrLeuProAsnThrValIleGlnThrLeuArgValSerGlyValTyrMetLeuLeu 140  
 Db 361 AGCTACTGCGCAACACTGTATTGAGACCCCTGCTGTCAGTGTGCAATGAGACTACTG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 421 TTGAGCCGAGTGGGAGACGACGCTGCTGCTACCTGCTGGACACATGCTCTTATCTT 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysVal 180  
 Db 481 CTGTGTGCCCCCAGCTGTGTGCTACCAAGGTGTGGGTCTCCCTGTACCAATTTGTCC 540  
 QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 541 ACCGCGATATCTGGCCCTCTGTGTCCGTTAGTACAGGCCACCCGACCCGTGGCAG 600  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleIleYsgSerSerSerArgGlnAlaPro 220  
 Db 601 AATTCACTAATCTTGTAGTCTTACCAACAGATCAAGACAGATGCGCAGAGACACCG 660  
 QY 221 LysProLeuAlaLeuProSerArgGlyTyrThrLysArgHisLeuSerLeuThrSerThr 240  
 Db 661 AAACCCCTGGCTGCTGCATCTCAGGATCAAGAGGACATGATCTCACAGTCAAGT 720  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGlnGlnLysProHis 260  
 Db 721 GTGCTTCACTAAGAGGACAGATCTATCTGTGCCAGAGTGAAGAGGAGGACCCAC 780  
 QY 261 ArgGlnValLeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerPro 280  
 Db 781 AGGAGGATGTACCAACCCCATGAGCAATATCATGGGTCCAAAGTCTGCTGCTCCCC 840  
 QY 281 GlnValProThrAlaGlnLysAspLeuSerSerTyrGlyLysValSerAspLeuSerLeu 300  
 Db 841 GAGGTGCTATCTCAGAGAAAGATTTGTCTTCAAGAAAGTGTGACCTGAGTCTC 900  
 QY 301 SerGlySerValCysCysLysHisLysProSerSerThrSerLeuLeuSerProArg 320  
 Db 901 TCTGGGTGCGTGTGCTGTAAACCAAGCCAGCTCCATCTCTGTGCAACACCCGCG 960  
 QY 321 GlnAsnAlaPheGlnLeuArgProPheIleGlnThrArgHisPheLeuTyrSerArgGly 340  
 Db 961 CAAATCCCTTTCAGCTCAGCCATTTATTTAGACCAAGCATTTCTTATCTCCAGGGGA 1020  
 QY 341 AspGlyGlnLysArgLeuLeuProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThr 360  
 Db 1021 GATGGCCAAAGCGTCTAAACCCCTCATCTCATGACCACTCCAGCTTAACTTGACT 1080  
 QY 361 GlyAlaArgArgLeuValGlnIleIlePheLeuGlySerArgProArgThrSerGlyPro 380  
 Db 1081 GGGGCCAGAGACTGGTGGAGATCATCTTCTGGGCTCAAGGCTTAAAGATCAAGACCA 1140  
 QY 381 LeuCysArgThrHisArgLeuSerArgArgTyrTyrGlnMetArgProLeuPheGln 400  
 Db 1141 CTCTGACAGACACACCGTCTATGCGCTCATCTGCAATGCAATGCGGCCCTGTTCACAG 1200  
 QY 401 LeuLeuValAsnHisAlaGlnCysGlnTyrValArgLeuLeuArgSerHisCysArgPhe 420

```
Db 1201 CTGCTGCTGAACCATGAGAGTCCCAATATGTCAGACTCTCTCAGGTGCATTCGACGTTT 1260
Qy ArgThrAlaAsnGlnGlnValThrAspAlaLeuAspThrSerProProHisLeuMetAsp 440
Db 1261 CGAACGACAAACCAACGAGTGAAGATGCTTGAAACACGACCCACCTCAGCTCAT 1320
Qy LeuLeuArgLeuHisSerSerProTyrPglValTyrGlyPheLeuArgAlaCysLeuCys 460
Db 1321 TTGCTCCGCTGCACAGACATCCCTGCGAGTATATGTTTCTTCGAGGCTGTCTTC 1380
Qy LysValValSerAlaSerLeuTyrGlyThrArgHisAsnGlnArgArgPhePheLysAsn 480
Db 1381 AAGGTGTGTCTGTAGTCTCTGGGATACAGGCAATATGACGCGCTTCTTAGAAC 1440
Qy LeuLysLysPheHisSerLeuGlyLysTyrGlyLysLeuSerLeuGlnLeuMetTyr 500
Db 1441 TTAAAGAGTTCACTCTGTGGGAAATACGGCAAGCTATCACTGAGAGACTGATGTGG 1500
Qy LysMetLysValGlnAspCysHisTyrPheLysSerSerProGlyLysAspArgValPro 520
Db 1501 AAGATGAAGTAGAGATTGCCACTGGCTCCGACAGCCCGGGAGAGACCGTGTCCC 1560
Qy AlaAlaGlnHisArgLeuArgGlnArgGlyLeuAlaThrPheLeuPheTyrLeuMetAsp 540
Db 1561 GCTGCAGAGACCGCTGAGAGAGAGATCTGGCTACGCTTCCTGCTTGGCTGATGAC 1620
Qy ThrTyrValValGlnLeuLeuArgSerPhePheTyrTyrGlnGlnSerThrPheGlnLys 560
Db 1621 ACATACCTGATCAGCTGCTTAGTTCATTTTATCATCAGAGACACATTCACAGAG 1680
Qy AsnArgLeuPhePheTyrArgLysSerValTyrPheSerLysLeuGlnSerTyrGlyValArg 580
Db 1681 AACAGGCTCTTCTCTACCTGTAAGAGTGTGAGAGACCTGACAGGACTTGATGATCAGG 1740
Qy GlnHisLeuGlnArgValArgLeuArgGlnLeuSerGlnGlnValArgHisGln 600
Db 1741 CAACACCTTGAAGAGTGGCGCTACGGAGCTGTCAACAGAGAGGTCAAGCATCACAG 1800
Qy AspThrTyrPheLeuAlaMetProIleCysArgLeuArgPheIleProLysProAsnGlyLeu 620
Db 1801 GACACCTGGCTAGCACTGCCATCTGCAAGCTGCTTCATCCCAAGCCCAAGCGGCTG 1860
Qy ArgProIleValAsnMetSerTyrSerMetGlyThrArgAlaLeuGlnArgArgLysGln 640
Db 1861 CGGCCCATGTGAACATGAGTATAGCATGGGTACACAGAGCTTTGGGCGAAGAGAGAG 1920
Qy AlaGlnHisPheThrGlnArgLeuLysThrLeuPheSerMetLeuAsnTyrGlnArgThr 660
Db 1921 GCCCAGCATTTCAACCGGCTCTCAAGACTCTTCAAGCATGTCTCAACTATGAGCGAGCA 1980
Qy LysHisProHisLeuMetGlySerSerValLeuGlyMetAsnAspIleTyrArgThrTyr 680
Db 1981 AATCATCTCACTCACTTAAGGGGTCTTCTGTACTGGGTATGATGACATCTCAACGAGCTGG 2040
Qy ArgAlaPheValLeuArgValArgAlaLeuAspGlnThrProArgMetTyrPheValLys 700
Db 2041 CGGSCCTTTGTGCTGTGTGCTGTGCTGTGACCAACACACCGAGATGATCTTGTTAG 2100
Qy AlaAspValThrGlyAlaTyrAspAlaIleProGlnGlyLysLeuValGlnValValAla 720
Db 2101 GCAGATGTGACCGGGGCTATGATGCGATCCCCAGGGTAAAGGTGTGAGGTTGTCC 2160
Qy AsnMetIleArgHisSerGlnSerThrTyrCysAlaArgGlnTyrAlaValArgArg 740
Db 2161 AATATGATCAAGCACTCGAGAGACACTACTGATATCCGCCAGATGACATGGGTGGAGAA 2220
Qy AspSerGlnGlyGlnValHisLysSerPheArgArgGlnValThrThrLeuSerAspLeu 760
Db 2221 GATAGCCAGGCGCAAGTCCCAAGTCTTAGGAGACAGGTCAACCTCTCTGACCTC 2280
Qy GlnProTyrMetGlyGlnPheLeuLysHisLeuGlnAspSerAspAlaSerAlaLeuArg 780
Db 2281 CAGCATATAGGGCCAGTTCCTTAAGCATCTGACAGATTCAATGACAGTGCAGTGAAG 2340
Qy AsnSerValValIleGlnGlnSerIleSerMetAsnGlnSerSerSerLeuPheAsp 800
Db 2341 AACTCGTGTATGATGACAGAGCATCTCTATGATGATGAGACACACACACGCTGTAGAC 2400
Qy PhePheLeuHisPheLeuArgHisSerValValLysIleGlyAspArgCysTyrThrGln 820
Db 2401 TTCTTCCTGCACTTCTCTGCTGACAGTGTCTGTAAGAATGGTGTACAGTGTATAGCAG 2460
Qy CysGlnGlyIleProGlnGlySerSerLeuSerThrLeuLeuCysSerLeuCysPheGly 840
Db 2461 TGCAGGGCATCCCCAGAGGCTCCAGGCTATCCACCCTGCTCTCAGTCTGTGTGGGA 2520
Qy AsnMetGlnAsnLysLeuPheAlaGlnValGlnArgAspGlyLeuLeuLeuArgPheVal 860
Db 2521 GACATGAGAAACAACTGTTTCTAGGTGACAGGAGATGGGTGCTTTACGTTTGT 2580
Qy AspAspPheLeuLeuValThrProHisLeuAspGlnAlaLysThrPheLeuSerThrLeu 880
Db 2581 GATGACTTCTCTGTGTGACGCTCATTTGACCAAGCAAAACCTTCTCAGACACCTG 2640
Qy ValHisGlyValProGlnTyrGlyCysMetIleAsnLeuGlnLysThrValValAsnPhe 900
Db 2641 GTCCATGGCGCTTCGATGATGGGTGATGATTAACCTTGACAAAGACAGTGTAACTTC 2700
Qy ProValGlnProGlyThrLeuGlyGlyAlaAspProTyrGlnLeuProAlaHisCysLeu 920
Db 2701 CCGTGGAGCCCTGATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Qy PheProTyrCysGlyLeuLeuLeuAspThrGlnThrLeuGlnValPheCysAspTyrSer 940
Db 2761 TTTCCCTGTGTGGCTGTGTGTGTGACACTGACCTTGGAGGGTGTCTGTGATCTTCA 2820
Qy GlyTyrAlaGlnThrSerIleLysThrSerLeuThrPheGlnSerValPheLysAlaGly 960
Db 2821 GGTATGCCCAACCTCAATTAAGACAGCTCACTTCCAGAGTGTCTCAAGCTCGG 2880
Qy LysThrMetArgAsnLysLeuLeuSerValLeuArgLeuLysCysHisGlyLeuPheLeu 980
Db 2881 AAGACCATCGGAAACAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
Qy AsnLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLysIlePheLeuLeu 1000
Db 2941 GACTTCAGATGAACAGCTCCAGACAGTGTGATCATATATACAGATCTTCTGCT 3000
Qy GlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArgValArgLys 1020
Db 3001 CAGGCTTACAGGTTCAATGATGATGTGATTCAGCTTCCCTTGAACAGCGGTAGAGAG 3060
Qy AsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysTyrAlaIleLeu 1040
Db 3061 AACCTCAATCTCTTCTGGGATCATCTCCAGCCAAAGATCTGCTGTATGATCTCTG 3120
Qy LysValLysAsnProGlyMetThrLeuLysAlaSerGlySerPheProProGlnAlaAla 1060
Db 3121 AAGCTGAAGAATCCAGGAATGACATTAAGGCTGTGCTCTTCCCTCGAAGCCGCA 3180
Qy HisTyrPheCysTyrGlnAlaPheLeuLeuLysLeuAlaHisSerValIleTyrLys 1080
Db 3181 CATTTGCTGTGCTTACCAAGCCCTTCTGCTCAAGCGTGTCTCATTTCTGATCTTAA 3240
Qy CysLeuLeuGlyProLeuArgThrAlaGlnLysLeuLeuCysArgLysLeuProGlnAla 1100
Db 3241 TGTCTCTGGACCTTGAAGACACCCCAAACGCTGTGTGCGGAGCTCCAGAGGCG 3300
Qy ThrMetThrIleLeuLysAlaAlaAspProAlaLeuSerThrAspPheGlnThrTyr 1120
Db 3301 ACATATACCATCTTAAAGTGCAGCTGACCCAGCCCTTAACACAGACTTTCAAGACATT 3360
Qy LeuAsp 1122
Db 3361 TTGGAC 3366
```

```

RESULT 2
AF051911      3426 bp  mRNA  linear  ROD 01-DEC-2000
LOCUS      Mus musculus telomerase reverse transcriptase mRNA, complete cds.
DEFINITION
ACCESSION  AF051911
VERSION     AF051911.1 GI:3005591
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 3426)
AUTHORS     Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
TITLE       Expression of mouse telomerase reverse transcriptase during
            development, differentiation and proliferation
JOURNAL     Oncogene 16 (13), 1723-1730 (1998)
MEDLINE     9821176
PUBMED      9582020
REFERENCE   2 (bases 1 to 3426)
AUTHORS     Greenberg, R.A., Allsopp, R.C., Chin, L., Morin, G.B. and DePinho, R.A.
TITLE       Direct Submission
JOURNAL     Submitted (02-MAR-1998) Microbiology and Immunology, Albert
            Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY
            10461, USA

FEATURES
source      1..3426
            Location/Qualifiers
            ..orgnism="Mus musculus"
            ..mol_type="mRNA"
            ..db_xref="taxon:10090"
            ..chromosome="13"
            ..map="WGD offset position 40; 40 cm from centromere"
            ..30..3398
            ..note="MTEXT"
            ..codon_start=1
            ..product="telomerase reverse transcriptase"
            ..protein_id="AAC09323.1"
            ..db_xref="GI:3005592"
            ..translacion="MTRAPRCAPVAVSLRREYRPPLATFRRLLPSSRRLLVQGGP
            KIRYLVAQCLVCMWGSQPPADLSHQVSLKELVAVVQRLCERNRNVLAFGE
            LNNARQGVPMVAFVSRYSLPNTVETLISGAMMLLSRVGDDLLVLLAHGALYL
            LVPSCAVQVGSPLVQICATTDIMPVSASYPTRPVGNFTNLFLQICSSROE
            APKPLAPSRGTKSLSTSTSVSAKAKCVPEVVEBPQVLPFTSGSWTSP
            ARSPVPAEKDLSKGVSDLSGSCVCKHKSSTLSPPROAFLRPFETRH
            FLXSRGQGERLNPSPLSNLQPLTARLVTILPESRPRSGPLCTHLSRYV
            QMRPLQGLVNAECQVRLNSHCRTFANQVTDALNTPHLMDLRLHSSPQ
            VYGLRACLCVVASLNGTRHNERPFRKULKFIISLGRGKSLQELMKRVDC
            WLRSPGKDRPAAHRLREILATFPLMDTVVQLASPFYITESTFQNRLEFY
            RKSVMGRALGRKQAHFORLKTPLSMANTERTPHLMSSVLGANDYRTWA
            NMSYMSKLOSIGVROHLERVALRELSOEVRHQDTMLAPICRLFIKPGNLPY
            FLVRLALDQTPRMVFKADVTGAYDAIPQGLVEVAVANIRSESTYCIROYAVVR
            DSGOVYHKSFRQYTTLSLOPYMGKPLKLDSDASALRNAYVIOSSIMSESSSL
            PDPFLHRLSHVYKIGRCCTCCGICQSSLSLCSLCFGMEKMLAEPORGGL
            LRFVDDFLVTFHLDQAKTFLSTVHGVPEYGCMLNQKTVNFVEPSTLGAAPYQ
            LPAHCLFPGQGLLDQTVLNVYKIFLQAVRPAACVQLPQPRVRLKTFELGIS
            LKCHGLFLLQVNSLQTVCLNIYKIFLQAVRPAACVQLPQPRVRLKTFELGIS
            SOASCVAIILKVNKPGMTLKASGSFPEPAHMLCYQAFLKLAHSHVYKCLGPIRT
            AQLCLRLPRAWTILKAAADPLSTDFQTLID"

BASE COUNT  746 a      979 c      907 g      794 t
ORIGIN
Alignment Scores:
Pred. No.:      1.67e-299      Length:      3426
Score:          5901.00      Matches:      1122
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:             10      Gaps:      0

US-09-042-460-2 (1-1122) x AF051911 (1-3426)
QY      1 MetThArgAlaProArgCysProAlaValaArgSerLeuArgSerArgTyrArgGlu 20
|||||

```

```

Db      30 ATGACCCGAGGCTCTGCTGCCCCGCGGCGAGCTCTGCTGCGACGCCATACCGGAG 89
QY      21 ValTTrpProLeuAlaThrPheValaArgArgLeuGlyProGluGlyArgArgLeuValGln 40
Db      90 GTGTGGCCGCTGGCAACCTTGTGTGGCGCCCTGGGGCCCGAGGGAGGGCGGCTGTGGCA 149
QY      41 ProGlyAspProIysIleTyrArgThrLeuValaGlnCysLeuValCysMetHisTrp 60
Db      150 CCGGGGAGCCGAGATCTACCGCACTTGGTTGGCCCATGCGCTGTGTGATGCATCGG 209
QY      61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyGlnLeu 80
Db      210 GGCTCACAGCTCCACCTCCGACCTTCTCCACAGGTCATCCCTTAAAGAGCTG 269
QY      81 ValAlaArgValaGlnArgLeuGlyCysGluArgAngGluArgAsnValaLeuAlaPheGly 100
Db      270 GTGGCCAGGTTGTGGAGAGACTGTGGAGCGCCACAGAGAAAGTGTGCTGGCTTGGC 329
QY      101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
Db      330 TTGGACCTGCTTACAGAGCCAGAGCGGCGCTCCCATGGCTTCACTAGTAGAGTGGGT 389
QY      121 SerTyrLeuProMetThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
Db      390 AGCTACTTGGCCCAACCTGTATTAGAACCTTGCTGTCAAGTGGATGATGCTACTG 449
QY      141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
Db      450 TTGAGCCAGATGGGACAGACCTGCTGCTGCTACTGCTGAGCACTGCTGCTTATCTT 509
QY      161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
Db      510 CTGGTGGCCCCCGAGCTGCTCTTACAGAGGTGTGGTCTCCCTGTACCAATTGTGGCC 569
QY      181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
Db      570 ACCAGGATATCTGGCCCTGTGTGCTGCTGCTTACAGGCCACCCAGCCCGTGGCAGG 629
QY      201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLeuSerSerSerArgGlnGluAlaPro 220
Db      630 AATTTCACCTAACCTTGGTGTCTTACCAAGATCAAGACAGAGAGTGGCAGAGACCG 689
QY      221 LysProLeuAlaLeuProSerArgGlyThrTyrAsnGlnHisLeuSerLeuThrSerThrSer 240
Db      690 AAACCCCTGGCTTGCATCTGAGGTAAAGAGGCAATGAGTCAACACAGTACAGT 749
QY      241 ValProSerAlaIysIysAlaArgCysTyrProValProArgValGluGlnGlyProHis 260
Db      750 GTGGCTTCACTTAAGAGAGCCAGATGCTATCTGTCCGAGGTGAGAGAGGAGCCAC 809
QY      261 ArgGlnValLeuProThrProSerArgIysSerTrpValProSerProAlaArgSerPro 280
Db      810 AGGCAAGGTGCTTACCAACCCCATCAGGCAATATGAGTGCCAAAGTCTGCTGCTGCC 869
QY      281 GluValProThrAlaGluIysAspLeuSerIysGlyIysValSerAspLeuSerLeu 300
Db      870 GAGGTGCTACTCAAGAGAAAGTTGTCTTAAAGAAAGTGTGACCTGAGTCTC 929
QY      301 SerGlySerValCysCysIysHisIysProSerSerThrSerLeuLeuSerProArg 320
Db      930 TCTGGGTGGGTGTGCTGTAACCAAGAGCCAGCTCACTCTGCTGCTACACCCGCC 989
QY      321 GlnAsnAlaPheGlnLeuArgProPheIleGluThrArgHisPheLeuTyrSerArgIly 340
Db      990 CAATATCCTTTAGCTCAGGCACTTATTAGACACAGCATTTCTTATCTCCAGGGGA 1049
QY      341 AspGlyGlnIysArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThr 360
Db      1050 GATGGCCAAAGGCTTAAACCCCTATCTTCACTCAGCAACCTCCAGCTTAACTGACT 1109
QY      361 GlyAlaArgArgLeuValGluIleIlePheLeuGlySerArgProArgThrSerGlyPro 380
Db      1110 GGGGCCAGAGAGACTGTGGAGATCATCTTCTGGGTGTAAGGCTTAGGATCATAGAGACA 1169

```



QY 381 Leucylsargthrhlsargleuserhargargytrtrnglmetarproleupheingln 400  
Db 1170 CTGTGACAGACACCCTCTATGCGCTGATGAGTGGAGTGGGCGCTGTTCACACG 1229  
QY 401 LeuleuValaenhlsalaglucysglntryValargleuleuargserhiscysargphe 420  
Db 1230 CTGCTGGTGAACATGACAGAGTGCATATGATGACAGCTCTGACGTGCATTCAGAGTT 1289  
QY 421 ArgthrhlaaenginglnValThrspalaleuanshrhserprohlsleumetasp 440  
Db 1290 CGAACAGCAACCAACAGGTGACGATGCTTGAACACCGCCACCGCATTCATGAT 1349  
QY 441 LeuleuargleuhserserprotrnglnValtryglpheleuarglaCysleucys 460  
Db 1350 TTGCTCCGCTGCACACAGTCCCTGGCAGGTATGTTTTCTTGCGGCTGTCTCTCC 1409  
QY 461 LysValValSerlaserleutrglyThrArgHlsanglularghPheleuysan 480  
Db 1410 AAGGTGGTGTCTGATGCTCTGGGGTACACAGGACACATGAGCGCGCTTCTTAAAGAAC 1469  
QY 481 LeulyslsPhehlseserleuGLysTryGLysleuserleuenglnleumettrp 500  
Db 1470 TTAAAGAGTTCATCTGTTGGGAAATACGGACACTTACATCCAGAACTGATGTGG 1529  
QY 501 LysMetLysValgluaspCysHlsrtrpleuargserprogllyysaspargValPro 520  
Db 1530 AAGATGAAGATGAGGATGTCCTGCTCCGACGACCGCGGGAGAGACCGTGTCCCC 1589  
QY 521 AlaAlaglunhlsargleuarglularglileuulathrPheleuPheleuMetasp 540  
Db 1590 GCTGACAGACACCGCTGAGGAGAGAGATCTGCTACGTTCTGTTCTGCTGATGAC 1649  
QY 541 ThrtryValValgluleuLeuargserPhepheytrlierhngluserThrPheglins 560  
Db 1650 ACATACGTGTACAGCTGCTTACGTTCATTTCTTACATCACAGAGACATTCACGAAG 1709  
QY 551 AsnargleuPhepheytryArgLysSerValTrpserLysleuenglnserlieglValarg 580  
Db 1710 AACAGGCTCTTCTCTACCGTGAAGAGTGTGAGCAAGCTGACAGCATTGAGTCAAG 1769  
QY 581 GlunhlsleuaglucysValargleuargluleuserglnglulValArgHlsngln 600  
Db 1770 CAACACCTTGAGAGAGTGGCTACGGAGCTGCACAAAGAGGTCAAGCATTCACG 1829  
QY 601 AsptrhrtrpleuulamecProleCysargleuargPheleprolyspProaenglyleu 620  
Db 1830 GACACCTGGTAGCCATCCCATCTGCAGACTGCGCTTCATCCCAAGCCCAACGCGCTG 1889  
QY 621 ArgprolleValaammetserTryserMetglYThrarglaleuaglYargaglysgln 640  
Db 1890 CGGCCATTTGACATGATGATTATGATGAGTGGTACCGAGCTTTGGCAGAGAGAGCG 1949  
QY 641 AlaGlunhlsPhehtrnglnargleuLysThrleuPheuserMetleuanshrygluargthr 660  
Db 1950 GCCCAGCATTCACCCACGCTCTCAAGCTCTTTCAGATGCTCAACTATGACGGACA 2009  
QY 661 LysHlsProhlsleumetglYserSerValleuclYmetasnpillettryargthrttrp 680  
Db 2010 AAACATCTCACTTATGGGTCTTCTGTACTGGGTATGATGACATTCACAGAGCTGG 2069  
QY 681 ArgAlaPheValleuargValarglaleuaspglnthrProargMetTryrPheVallys 700  
Db 2070 CGGGCTTTGTGCTGCGTGTGCTGCTGACACAGACACCGAGAGTACTTGTGTAAG 2129  
QY 701 AlaaspValThrnglyValtryaspalalleproglnglyLysleuValgluValala 720  
Db 2130 GCAGATGTGACCGGGGCTATGATGCCATCCCGAGGGTAAAGGTGGAGGTGTGTTGCC 2189  
QY 721 AsmetlleargHlsesergluserThrTryCyslleargglntryalaValalaarg 740  
Db 2190 AATATGATCAGGCACTCGGAGAGACGTACTGTATCCCGCAGTATGAGTGTCCGAGAG 2249

QY 741 AspserGlnlyGlnValHlsYserSerPheargarglnValThrtrhsPheleu 760  
Db 2250 GATAGCAAGAGCCAGATCCACAAAGCTTTTAGAGACAGGTCCACACCTCTGACCTC 2309  
QY 761 GlnProtryMetglYGlnPheleuLysHlsleuGlnaspSeraspAlaSerAlaLeuarg 780  
Db 2310 CAGCCATTAACATGGGCGAGTCTTATAGCATTCGACAGATTAAGATGCCAGTGCAG 2369  
QY 781 AsnSerValValilleglnglnserlieSerMetCaengluserSerSerleuPheasp 800  
Db 2370 AACTTCCTGTATCATCAGACAGACATCTTATGATGAGAGAGAGAGAGCTGTTTAC 2429  
QY 801 PhePheleuHlsPheleuargHlsSerValVallysllieglYaspargCysTryThrGln 820  
Db 2430 TTCTTCCTGCACTTCTGCGTGCACAGTGTCTGAAGATTTGGTGCAGGTGTCTATACGAG 2489  
QY 821 CysGlnGlylileproglnglySerSerleuserThrleuLysSerleuCysPhegly 840  
Db 2490 TGCAGGGCATTCGCCAGGAGCTCCAGCTCCAGCTTACCTCTGCTGAGTGTGCGA 2549  
QY 841 AspMetGlnasnlYleuPhealaglValGlnargaspGlyleuLeuLysPheVal 860  
Db 2550 GACATGAGAAACAAGCTGTTGCTGAGTGCAGCGGATGGGTTCCTTTACGTTTGT 2609  
QY 861 AsnAspPheleuLeuValThrProhlsleuaspGlnAlalystrPheleuserThrleu 880  
Db 2610 GATGACTTCTGTGTGTGACGCTCATCTTGACCAAGCAAAACCTTCTCAGACCTG 2669  
QY 881 ValHlsGlyValProgluntryglYCysMetlieasnlleuGlnLysThrValalaanphe 900  
Db 2670 GTCCATGCGCTTCTAGTATGGGTGCAGATGAATTCGACAAAGAGTGTGAATTC 2729  
QY 901 ProvalGlnProglYThrleuGlyValalaaprotryGlnleuProAlaHlsCysleu 920  
Db 2730 CCGTGGAGCGCTGTACCTGGGTGTGAGCTCCATACACCTCCCTGCTCACGCTG 2789  
QY 921 PheProtrpCysglYleuLeuLeuaspThrnglnThrleuGlnValPheCysaspTrySer 940  
Db 2790 TTTCCTGTGTGTGCTGTCTGTGACATCCAGACTTGTGAGAGTGTCTGTACATCTCA 2849  
QY 941 GlyTryAlaGlnThrserlieLysThrserleuThrPheGlnSerValPheLysAlaGly 960  
Db 2850 GGTATCCCAAGCTCAATTAAGAGAGACCTCACCTTCAAGGTGTCTTCAAACTGGG 2909  
QY 961 LysThrMetArgasnlYleuLeuSerValleuargLysCysHlsGlyleuPheleu 980  
Db 2910 AAGACATGCGGAACAAGCTCTGTGTGTGCTGTGAGTGTCAAGGTATTTCTTA 2969  
QY 981 AspLeuGlnValaanserleuGlnThrValCyslleasnlletryLysllePheleuLeu 1000  
Db 2970 GACTTGAGGTGAACAAGCTCCACAGAGCTGCATCAATATATACAAAGATCTTCTGCTT 3029  
QY 1001 GlnAlaTryArgPheHlsalacysValilleGlnleuProPheaspGlnargValarglys 1020  
Db 3030 CAGGCTTACAGGTTCATGATGTGATTCAGCTTCCCTTTCACAGCGGTATGAGAG 3089  
QY 1021 AsnleuThrPhePheleuGlyllelleserSerGlnAlaserCysTryValalleu 1040  
Db 3090 AACTTCATCTTCTTGTGGCACAATCTCAGCAAGACATCTGCTGCTATGCTATCTG 3149  
QY 1041 LysVallylsanProglYmetThrleuLysalaserGlySerPheProGlnulala 1060  
Db 3150 AAGGTCAAGATCCAGAAATGACACTAAAGGCTCTGCTCTTCTCTCCAGACCGCA 3209  
QY 1061 HlsrtrpleuCysTryGlnAlaPheleuLysleuAlaHlseserValilletryLys 1080  
Db 3210 CATGGCTTGTCTACAGGCTTCTGCTCAAGCTGCTGCTCATTTCTGTATCTACAA 3269  
QY 1081 CysleuLeuGlyProleuargThrAlaGlnLysleuLeuCysargGlyleuProGlnula 1100  
Db 3270 TGTCTCTGGACCTGTGAGAGACAGCCAAAACCTGCTGTGCGGAAGTCCCAAGGCG 3329  
QY 1101 ThrmetThrilleuLysAlaAlaaspProAlaLeuserThrAspPheGlnThrille 1120

DB 3330 AACATGACCATCTTAAAGCTGACGACCCAGCCCTTAAAGACAGCTTCAGACCATTT 3389  
QY 1121 LeuASP 1122  
DB 3390 TTGGAC 3395  
RESULT 3  
AF149012 4170 bp mRNA linear ROD 21-FEB-2001  
LOCUS Mesocricetus auratus telomerase catalytic subunit mRNA, complete  
DEFINITION cds.  
ACCESSION AF149012  
VERSION AF149012.1 GI:6572639  
KEYWORDS  
SOURCE Mesocricetus auratus (golden Syrian hamster)  
ORGANISM Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.  
REFERENCE 1 (bases 1 to 4170)  
AUTHORS Guo, M., Okamoto, M., Lee, Y. M., Baluda, M. A. and Park, N. H.  
TITLE Enhanced activity of cloned hamster TERT gene promoter in  
transformed cells  
JOURNAL Biochim. Acta 1517 (3), 398-409 (2001)  
MEDLINE 21240330  
PUBMED 11342218  
REFERENCE 2 (bases 1 to 4170)  
AUTHORS Guo, M., Okamoto, M., and Park, N. H.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAY-1999) Dentistry, University of California, Los  
Angeles, 10833 Le Conte Ave., Los Angeles, CA 90095, USA  
Location/Qualifiers  
1. 4170  
/organism="Mesocricetus auratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10036"  
/note="common: golden Syrian hamster"  
28. 3414  
/note="hamTERT"  
/product="telomerase catalytic subunit"  
/codon\_start=1  
/protein\_id="AAF17334.1"  
/db\_xref="GI:6572640"  
/translation="MPRAPRCRAVALLISGYRQVWPLATVRELGEGHQVQDP  
KVERIVARCLVCPMDSPPADLSFQVSLKEIVARVQRCERGENVATFGA  
LNGAGCGPMTFTTSYSLPNSVTESLRYSGAMMLINNVQDLVYLINCAVL  
LVPPSCAYOVGSPYQICATAEIWPVSRIYPRPGVGNFTLSTHERVNSHOB  
AKKPPPLREKAKRSISTNRSVSPSKKACDLPRLKGPVQAVTPSDKTVPP  
AKSHAVPIRTTKEDLSGVKAPGLSRGVCYKQSPSTLSQPLCONA.POLRPT  
TRKPLVSRGGERLNPSTLNPLOPSITGABRLVEITFLMRPTSGPICGRRLSK  
RTWQMPPLFQQLVNHARCVPYRLRSLRCPRTAAHQVAGLNTISFORLMLRLHS  
SPQVQVFLQACVGLKVPGLGSRHNRREFKMKVERISLGRYDKSLDELTWQKV  
QDCRMRLSSPGNNVCVAEHRTRERILAVFLFIMDAVYVILRSPVYETFKM  
LEFYKSMWRLOSIGVRLHREVRLOELSEVROEAMPAPICRLPIPKPSGL  
RATVYNSWVGTRAFPGKQAOHFTQCLTISVNYELTKTNLIGASVGLNDIYT  
WRRTVARTLDPAPMYEVADYTGAVDAIPQDKLVITAMTRHDNSGCHQVAV  
VQDRQGLHKFRQVSTLSDLPQMGQDLKHODBSTLSLNSVITBESLSNBS  
SSLDFELFRVNSVVKIGRCYVQCGIPQSSLSLTLCSLCEGDMENKLFARVQDD  
GLILRFVDDPLLVTPHLVQAEFLALVAGIPEYGMINLQKTVNPEVAGITDGA  
PHQLPALCFPMGGLDQLEVLCDYTGARTSIKASLTFORFAGRMQRKLA  
VLTKCHSLFLDKMNSLQTCINIVYKIFLOAYFHAQLQDPDQVHRNPPAFIS  
LISNASCCTSLTKVNNAGMTKAGAGSGSPPEARWLYQAFLLKLAGHSVYKCL  
LGPLRTRQQLCRKLPRAIMATLIEPAAADPALSTDPQTLID"

BASE COUNT 927 a 1177 c 1126 g 940 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3,15e-239 Length: 4170  
Score: 4751.00 Matches: 904  
Percent Similarity: 88.04% Conservative: 90  
Best Local Similarity: 80.07% Mismatches: 127  
Query Match: 80.01% Indels: 8

DB: 10 Gaps: 4  
US-09-042-460-2 (1-1122) x AF149012 (1-4170)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
DB 28 ATGCCCGCGCGCGCGCGTTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 87  
QY 21 ValThrProLeuAlaThrPheValArgArgLeuGlyProGluGlyValArgArgValGln 40  
DB 88 GTGTGGCG 147  
QY 41 ProGlyAspProIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
DB 148 CCGCGGAGCCAAAGGCTTCGCCGAGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 207  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80  
DB 208 GACTCACAACTCCACTGCTGACCTTCTTCACACAGGTGTCATCTGAAGAGACTG 267  
QY 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
DB 268 GTGCCCGAGGCTCGCGAGACTCTGCGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 327  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
DB 328 TTCCGCGCTTAAACGAGCGCAAGGCGGCTCCCATGACATTCACACAGCGTGGCGC 387  
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaThrMetLeuLeu 140  
DB 388 AGCTRCCGCCCACTCGGTGACTGACTCTCTCGGTGACTGCTGCTGCTGCTGCTGCTGCTG 447  
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
DB 448 CTGAACCGAGTGGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507  
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
DB 508 CTGGTGGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 567  
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
DB 568 ACCGAGAAACCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 627  
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLeuSerSerSerArgGlnGluAlaPro 220  
DB 628 AATTTTACTCATCTTGATTCACACACCGGCTGAGACAGCTGACCGAGAGCATGG 687  
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
DB 688 AAACCCCGCGCGCTTGCATCTCGATCCAGAGCGGAGGAGTCTAAGCATCACCAATAGAGT 747  
QY 241 ValProSerAlaValGlyAlaArgCysTyrProValProArgValGluGluGlyProHis 260  
DB 748 GTGCTTCATCTTAAGAGAGCGGAGTGGATCTGGCGCGAGCTGAGAGAGGAGACCTTAC 807  
QY 261 ArgGlnValLeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerPro 280  
DB 808 AGCGAGGAGATTCACACCCCATCAGACAAACAGGGTGCAAAATCTGCAAGTCCCAT 867  
QY 281 GluValPro-----ThrAlaGluLysAspLeuSerSerLysGlyValSerAsp 297  
DB 868 GCAGTGCCTATTAGTAACTACCAAGAGATTTGCTTCGCGGAGTGAAGGACCTGGC 927  
QY 298 LeuSerLeuSerGlySerValCysCysLeuHisIleYspProSerSerThrSerLeuLeuSer 317  
DB 928 CTGAGTGCCTCTGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 987  
QY 318 ProProArgGlnAsnAlaPheGlnLeuArgProPheIleGluThrArgHisPheLeuTyr 337  
DB 988 CCACGTGTGCAAAATGCTTCACTGACACTATATCTAGACCAAAACCTTCTCTTAC 1047  
QY 338 SerArgGlyAspGlyGlnGluArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnPro 357

```

Db      1048  TCTAGGAGAGGTGGCCGAGAGAGGCTGACCCCTCGTTCTTACTCAACCACTGACGCC 1107
Qy      358  AsnLeuThrGlyAlaArgArgLeuValGluIlePheLeuGlySerArgProArgThr 377
Db      1108  AGCTTGACTGGGGCCAGAGAGACTGTGAGAGACTCTTTCTAGGCACTGAGACTTACGACA 1167
Qy      378  SerGlyProLeuGlyArgThrHisArgLeuSerArgThrTyrTrpGlnMetArgProLeu 397
Db      1168  TCGGACCACTCTGTGGAGACGCCGCTATCGAAGGCTACTGGCAATGGCGGCCCCCTTA 1227
Qy      398  PheGlnGlnLeuLeuValAsnHisAlaGlnCysGlnTyrValArgLeuLeuArgSerHis 417
Db      1228  TTCCAGCAGTTGCTGTGACCATCGCGGGGCCGATGTCCGGCTCCAGTCCCAT 1287
Qy      418  CysArgPheArgThrAlaAsnGlnGlnValThrAspAlaLeuAsn---ThrSerProPro 436
Db      1288  TGCAGGTTTCGACCGGACGCCACACAGGTGGCAGGTCCCTGAACACACACACGCCACAG 1347
Qy      437  HisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyrGlyPheLeuArg 456
Db      1348  CGCCTCATGAATTGCTCCGTCTACAGCAGAGTCCCTGGCAGGTAATGGCTTTCTTACG 1407
Qy      457  AlaCysLeuGlyValValSerAlaSerLeuTrpGlyThrArgHisAsnGlnArgArg 476
Db      1408  GCCTGTGTGGAAAGCTGTGCTCCGGGTCTCTGGGGTTCCCGGCAACACCGCAGCC 1467
Qy      477  PhePheLeuAsnLeuLeuValPhePheIleSerLeuGlyLysTyrGlyLysLeuSerLeuGln 496
Db      1468  TTCTTTAAGAACGTGAACGGTCTATCTCCCTGGGGAAGTATGACAAAGCTGTGGCTGAG 1527
Qy      497  GlnLeuMetTrpLysMetLysValGlnAspCysHisTrpLeuArgSerSerProGlyLys 516
Db      1528  GAGCTGACGTGGAGATGAAGTTCAAGACTGCAGTGGCTCCGACGACGCCAGGGAAC 1587
Qy      517  AspArgValProAlaAlaGlnHisArgLeuArgGlnArgIleLeuAlaThrPheLeuPhe 536
Db      1588  AACTGTGTCCGGGTGCAAGACCGGACCGAGGAAAGAGATCCGTGTGTGTGTGTC 1647
Qy      537  TrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyrIleThrGlnSer 556
Db      1648  TGGCTGATGGACCGCTAGCTGTGAGCTGTGGTCACTTCTTTACGTCAACAGAGAC 1707
Qy      557  ThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSerLysLeuGlnSer 576
Db      1708  ACTTTCAGAAAGAACCGGCTCTTCTTACCGAAAGAGCATGTGGAGAGCGTCGACAGC 1767
Qy      577  IleGlyValArgGlnHisLeuGlnArgValArgLeuArgGlnLeuSerGlnGlnGlnVal 596
Db      1768  ATTGAGTCAAGCACTTGTGAGAGGTGGCTTCAAGAACTGTCTCAAGAAAGATC 1827
Qy      597  ArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArgPheIleProLys 616
Db      1828  AGGCAGCGCCAGAGAGCCCTGGCAGGCCATGCCCATCTGCAGACGCTTTTCACTCCCAAG 1887
Qy      617  ProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThrArgAlaLeuGly 636
Db      1888  CCCAGTGGCTTCCGCCCATTTGTGAACATGAGTTAT---ATGGCAACCAAGCCTTTGAC 1944
Qy      637  ArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPheSerMetLeuAsn 656
Db      1945  AAAGGGAAGCAGGCTCAGCATTTTCAACCAAGTGTCAAGACTGTTCACAGCTGTCAAC 2004
Qy      657  TyrGlnArgThrLysHisProHisLeuMetGlySerSerValLeuGlnMetAsnAspIle 676
Db      2005  TATGAACCTGACAAACATCTACTCTGGGGGATGTGATCTGGGCTCGAATGATATC 2064
Qy      677  TyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGlnThrProArgMet 696
Db      2065  TACAGGACTGGCGAGCTTGCTACTGTGTGGGCACTGTGACCCAGACCCAGAGATG 2124
Qy      697  TyrPheValLysAlaAspValThrGlyAlaTyrPheAlaIleProGlnGlnLysLeuVal 716

```

---

```

Db      2125  TACTTTGTAAAGCAGATGTGACAGGGGATATGATGCAATCCCCCAGGACAGCTGTG 2184
Qy      717  GluValValAlaAsnMetIleArgHisSerGlnSerThrTyrCysIleArgGlnTyrAla 736
Db      2185  GAGGTATTTGCCAATATGATGATCAACACCCAGCAACTGCTGTACTGTATCAACCAATATCA 2244
Qy      737  ValValArgArgAspSerGlnGlnValHisLysSerPheArgArgGlnValThrThr 756
Db      2245  GTGGTCAAGAGATATGACAGAGCCAAATCCCAAGTCTTCAAGAGACAGGTTCTCCAC 2304
Qy      757  LeuSerAspLeuGlnProGlyThrMetGlyGlnPheLeuValHisLeuGlnAspSerAla 776
Db      2305  CTCCTGACCTCCAGGCACACAGAGGCGAGTTCTTGAAACATCTTCAGGACTCAACACC 2364
Qy      777  SerAlaLeuArgAsnSerValValIleGlnGlnSerIleSerMetAsnGlnSerSerSer 796
Db      2365  AGTCGGCTGAGAACTCCGTGTGATGACAGAGTTTCTCTGAACAGAGCCAGCAGCC 2424
Qy      797  SerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLysIleGlyAspArg 816
Db      2425  AGCCTGTTCACATTCTTCCGCGCTTGTGCGTAACAGTGTGTGAAGATCGGTGGCAGG 2484
Qy      817  CysThrThrGlnCysGlnGlyIleProGlnGlnLysSerLeuSerThrLeuLeuCysSer 836
Db      2485  TGTATGTCCAGTGTCCAGGCAATCCCGCAGGGGTCCAGGCTGTCCACCTGTCTGCACT 2544
Qy      837  LeuCysPheGlyAspMetGlnAsnLysLeuPheAlaGlnValAlaArgAspGlyLeuLeu 856
Db      2545  CTGTGTTCCGGGACATGAGAGAACAGTGTGTGCTGAAGTCAACAGAGATGGCTGCTT 2604
Qy      857  LeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGlnAlaLysThrPhe 876
Db      2605  TTGGTTTGTATGATGCTTCTGTGTGGAGAACCTCACTGTGTCCAGGCGGAACCTTC 2664
Qy      877  LeuSerThrLeuValHisGlyValProGlnTyrGlyCysMetIleAsnLeuGlnLysThr 896
Db      2665  CTCAGGCGCTGTCCGTGGACACCTCTGAGTACGGGTGATGATTAACCTGCAAGAAC 2724
Qy      897  ValValAsnPheProValGlnProGlyThrLeuGlyValAlaAlaProGlnLeuPro 916
Db      2725  GTGGTAACCTTCCCTGGAGCGCTGTATCCCTGAAGGACAGCTCCACACACAGCTGCTT 2784
Qy      917  AlaHisCysLeuPheProTrpCysGlyLysLeuLeuAspThrGlnThrLeuGlnValPhe 936
Db      2785  GCTACGTGCTGTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2844
Qy      937  CysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThrPheGlnSerVal 956
Db      2845  TGTGACTTACCTGGTATGCTGGGACCTCAATTAAAGCCAGCCTCACTTCCAGCGCAC 2904
Qy      957  PheValAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArgLeuLysCysHis 976
Db      2905  TTCAAGCGGGAGAGAACATGCACAGAAAGCTTTAGCTGTTTGGGTGTAAGTGTAC 2964
Qy      977  GlnLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLys 996
Db      2965  AGTCTGTTTCTAGACTTGCAGATGATAGCTTCAACAAGTGTGATCATATGATGACAG 3024
Qy      997  IlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGln 1016
Db      3025  ATCTTCCTGTCTCAGGCTCAGCTTCAAGTTCATAGCTGTGGCTCAGCTTCTTGAACCA 3084
Qy      1017  ArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysCys 1036
Db      3085  CATGTTGAAGAAAGACCCCGCATCTTCTGAGATCATCTCCAACTACACATCCGTGTC 3144
Qy      1037  TyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLys-----AlaSerGly 1053
Db      3145  TACTCCATCTCGAAGGTCAGAAATGACAGATACACTAAAGGCCAAGGCTCCTGTGCG 3204
Qy      1054  SerPheProProGlnAlaAlaHisTrpLysCysTyrGlnAlaPheLeuLysLysVal 1073
Db      3205  TCAATTCCTCTGAAGCTGACAGTGGCTGTGTTACCAAGCTTCTGTCTCAAGCTGGCT 3264

```

QY 1074 AAlhisservAlilleTyTyLeCysLeuLeuGlyProLeuArgThraAlaGlnLysLeuLeu 1093  
 DB 3265 GATCATTTCTGACCTACAGAGTGTCTCTGGAGCCTTCAGAGACAGACAAAACAGCTG 3324  
 QY 1094 CysArgLysLeuProGluAlaThraMetThrLleLeuLysAlaAlaAAspProAlaLeu 1113  
 DB 3325 TGCCGAAGCTCTCCAGAGGCAACATGGCCATCTTAAAGACTGCCCTGACCCAGCCCTTA 3384  
 QY 1114 SerThrAspPheGlnThrLleLeuAsp 1122  
 DB 3385 AGACAGACTTTCAGACCATTTTGAC 3411  
 RESULT 4  
 BD091553  
 LOCUS Adult bone marrow-origin cell capable of differentiating into 3396 bp DNA linear PAR 27-AUG-2002  
 DEFINITION myocardial cell.  
 ACCESSION BD091553  
 VERSION BD091553.1 GI:22637164  
 KEYWORDS WO 0148149-A/16  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 3396)  
 AUTHORS Umezawa,A., Hata,J., Fukuda,K., Satoshi, Ogawa and Sakurada,K.  
 TITLE Adult bone marrow-origin cell capable of differentiating into myocardial cell  
 JOURNAL Patent: WO 0148149-A 16 05-JUL-2001.  
 KYOWA HAKKO KOGYO CO LTD,AKIHRO UMEZAWA,TUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA,KAZUHIRO SAKURADA  
 COMMENT OS Homo sapiens (human)  
 PN WO 0148149-A/16  
 PD 05-JUL-2001  
 PF 28-FEB-2000 WO 2000JP001148  
 PR 28-DEC-1999 JP 99P 372826  
 PI AKIHRO UMEZAWA,TUNICHI HATA,KEIICHI FUKUDA,SATOSHI PI  
 OGAWA,KAZUHIRO SAKURADA  
 PC C12N5/06,C12N5/00,A61K35/28,A61P41/00,A61K48/00,C07K16/18 CC  
 FH Key Location/Qualifiers  
 FT CDS (1)-(3399).  
 FEATURES  
 source Location/Qualifiers  
 1..3396  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 549 a 1157 c 1088 g 602 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,62e-174 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: 6 Gaps: 13  
 US-09-042-460-2 (1-1122) x BD091553 (1-3396)  
 QY 1 MetThrArgAlaProArgCysProAlaValAlaArgSerLeuLeuArgSerArgTyArgGlu 20  
 DB 1 ATGCCGCGCGCTCCCGGCTGCCAGCCGCGGCTCCCTGCTGCCAGCCTACCGGAG 60  
 QY 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgGluValGln 40  
 DB 61 GTGCTGCCGCTGCCAGCTTCGTGGCGGCTGGGGCCCAAGGCTGGCGCTGGTGCAG 120  
 QY 41 ProGlyAspProLysLleTyArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 DB 121 CGCGGGAGACCCGCGGCTTCCGCGGCGCTGGCCAGATGCTGTGTGCTGCCCTG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGlnLeu 80

DB 181 GAGCAGCGCCGCCCGCCCGCCCTTCTCCCGCAGGTCTCTGCTGAGAGAGCTG 240  
 QY 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValAlaPheGly 100  
 DB 241 GTGGCCCGAAGTGTGAGAGAGCTGTGCGAGCCCGGCGGCGGCAAGAGTGTGCTTGCAGC 300  
 QY 101 PheGlnLeuLeuAsnGlnAlaArgGlyLysProPomeAlaPheThrSerSerValArg 120  
 DB 301 TTGCGCTGTGAGACGGGCGCCCGGGGGCGCCCGGAGCTTACACACAGCTTGGCG 360  
 QY 121 SerTyLeuProAsnThrValIleGlnThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 DB 361 AGTACTGCTCCCAACACGGTGAACGAGCATGTGCGGGAGCGGGCGGCTGCTGTG 420  
 QY 141 LeuSerArgValAlaAspAspLeuLeuValTyTrpLeuAlaHisCysAlaLeuTyTrp 160  
 DB 421 CTGGCGCGGTGGGCGACAGACGTGTGTACCTGTGGACAGCTGGAGCTTCTTGTG 480  
 QY 161 LeuValProProSerCysAlaTyArgValCysGlySerProLeuTyGlnLleCysAla 180  
 DB 481 CTGGTGGCTCCAGCTGGCGCTACAGAGTGTGGGGCGCGCTGACAGCTCGCGCT 540  
 QY 181 ThrThrAspLleTrpProSerValSerAlaSerTyArgProThrArgProValGlyArg 200  
 DB 541 GCCACTAGCGCCGCGCCCGCCAGCAGCTAGT--GGAACCCGAGAGCGCTTGGGA-- 594  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnLleLysSerSerSerArgGlnGlnAlaPro 220  
 DB 595 -----TCCGAACCGGCTGGAAACATAGGTACAGGAGCGCGG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
 DB 634 GTCCCGCTGGCGCTGCCAGCGCCGGGTGGAGAGGCGCGGGGCAAGTCCAGCGCAAGT 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyProValProArgValGluGluLysPro-- 259  
 DB 694 CTGGCGTTGCCAAGAGGCCAGCGGTGGCGCTGCCCTTGAAGCGGAGCGAGCGCCGTT 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp-- 272  
 DB 754 GGGCAGGGGTCTGGGCGCCAGCGGGCAGGACCGGTGAGCCAGATGACCGGTGTTCTGT 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerSerLys 292  
 DB 814 GTGGTGCACCTGCAGA-----CCGCGGAAGAGACCACTTTTGGAG 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer--GlySerValCysCysLysHisLysProSer 311  
 DB 859 GGTGCGCTCTGGCAGCGCCCATCTCCACCATCCGTGGGCGGACAGCAGCAGCGGGGC 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro--PheIle 330  
 DB 919 CCCCATTCACATCGCGCGCCAGCGCTCTGGAGACAGCCCTTGTCCCGGATGAGCGC 978  
 QY 331 GlnThrArgHisPheLeuTyTrpSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 979 GAGACCAAGACATTCCTACTCTCTCAGGAGC--AAGAGAGACGTGGCGGCTCTCTTC 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 DB 1036 CTACTCAGCTCTCGAGCGCCAGCGCTGAGCGGTGGAGGCTGTGAGACCACTTT 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
 DB 1096 CTGGCTTCAGGCGCTTGAATGCCAGGAGACTCCCGCAGATGTGCCCGCTGCCAGCGCC 1155  
 QY 391 TyTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTy 410  
 DB 1156 TACTGGCAATGGCGCCCTGTTCTGAGACTCTTGGGAACCAACCGGAGCGCCCTTAC 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430

Db 1216 GGGGTGCTCTCAAGAGCACTGCCGCTGGAGCTGCG-----GTCACCCCAAGCA 1266  
Qy 431 Leu-----Aen 432  
Db 1267 GCCGGTCTGTGCCCCGAGAGACCAGGGCTCTGTGGCGGCCCCCAAGAGAGAGAGCA 1326  
Qy 433 ThSerProProH1sLeuMetAspLeuLeuArgLeuHisSerSerProTropJinValTyr 452  
Db 1327 ACAGACCCCGCTGCGCTGTGAGCTCTCCGACAGACAGAGCCCTCGCAGCGTGC 1386  
Qy 453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTropJinThrArgHis 472  
Db 1387 GGCCTCGTGGCGGCTGTGCTGGCGCTGGTCCCAAGCCCTCGGGCTCCAGAGCAC 1446  
Qy 473 AengIuArgArgPhePheAsnLeuLeuValSerPheIleSerLeuGlyTyrGlyLys 492  
Db 1447 AACGAAACCCGCTCTCTAGAGAACCAACGAAAGTTCACTCTCCGGAGAACATGCCAAG 1506  
Qy 493 LeuSerLeuGlnGluLeuMetTropLysMetLysValGluAspCysHisTropLeuArgSer 512  
Db 1507 CTCTGCTGCAGAGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTGGCCAGG 1566  
Qy 513 SerProGlyLysAspArgValProAlaAlaGlnHisArgLeuArgGluArgIleLeuAla 532  
Db 1567 AGCCAGGGGCTTGCTGTGTCGCGAGAGCAACGCTGTGGTGAAGATCTGAGC 1626  
Qy 533 ThrPheLeuPheTropLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1627 AAGTCTGCACTGGCTGATGAGTGTGAGCGTGTGAGCTGCAAGCTTTCTTTAT 1686  
Qy 553 IleThrGlnSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTropSer 572  
Db 1687 GTACAGAGACACACGTTTCAAAAGACAGGCTCTTTTCAACCGAAGATGTCTGAGAC 1746  
Qy 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGlnArgValArgLeuArgGluLeuSer 592  
Db 1747 AAGTTCAAAGCATTTGAAATGACACACCTTGAAGAGGTGACGCTGGAGAGCTGTG 1806  
Qy 593 GlnGlnGluValArgHisHisGlnAspThrTropLeuLysMetProIleCysArgLeuArg 612  
Db 1807 GAAGCAGAGGTGCAGGCACTGCGGAAGCCAGGCGCCCTGCTGACGTCCAGACTGCCC 1866  
Qy 613 PheIleProLysProAsnGlyLeuArgTropIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1867 TTCACTCCCAAGCTGACGGGCTGGCGCATGTGTGAACATGACATCGTCGAGGAGCC 1926  
Qy 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1927 AGAAGCTCCGACAGAAAGAGGGCCGAGCGTCTCACCTGAGGGTGAAGCACTGTTTC 1986  
Qy 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 1987 AGCGTGCACACTGACAGACGGGCGCGCGCGCCCTCTGAGCGCGCTGTGCTGGGCG 2046  
Qy 673 MetAsnAspIleTyrArgThrTropArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2047 CTGACGATATCCACAGGCGCTGGCGACCTGTGCTGTGCTGTGGCGCCAGAGCCG 2106  
Qy 693 ThrProArgMetTyrPheValLysValAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2107 CCGCTGAGCTGTACTTGTTCAGAGTGAATGTGACGGGCGCGTACACACCACTCCCGAG 2166  
Qy 713 GlyLysLeuValGluValAlaValAsnMetIleArgHisSerGlnSerThrTyrCysIle 732  
Db 2167 GACAGGCTCAAGAGGTATATGCCACATCAAA---CCCAAGAACAGTACTGCGG 2223  
Qy 733 ArgGlnTyrAlaValAlaArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
Db 2224 CGTGGTATGCCGAGGTTCAGAGAGCGCGCCATGGGCACTCCGCAAGGCTTCAAGAGC 2283  
Qy 753 GlnValThrThrLysSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2284 CACGCTCTACCTTGAACAGCACTCCAGCCGTATCATGCAAGTTCTGTGGCTCACTGCA 2343

Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGlnGlnSerIleSerMetAsn 792  
Db 2344 GAGAC-----AGCCCGTGAAGAGTGCCTGTATCTGACACAGAGCTCTCTCTGAA 2397  
Qy 793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
Db 2398 GAGGCCAGCAGTGGCTCTTCAACGCTTCAACGCTTCAATGCTATGTCACACAGCCGTGCG 2457  
Qy 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
Db 2458 ATCAGGGCGCAAGTCTTACGTCAGTCCAGGGGATCCCGAGGCGCTTCACTCTTCCAG 2517  
Qy 833 LeuLeuCysSerLeuLeuPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
Db 2518 CTGCTCTGACGCTGTGTGATACGCGACATCGAAGACAGCTTTGGGGGATTTGGGCG 2577  
Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
Db 2578 GACGGGCTGCTCTGCGTGTGGTGAATGATTTCTGTGTGACACCTCACCTCACCCAC 2637  
Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGlyTyrGlyCysMetIleAsn 892  
Db 2638 GCGAAACCTTCTCTACAGACCTGATCCAGAGTGTCCATGATATGCTGCTGGTGAAC 2697  
Qy 893 LeuGlnLysThrValAlaAsnProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
Db 2698 TTGGGAAAGACAGTGTGAATCTTCCGTGAAAGACAGAGCCCTGGGTGGCAGGCTTT 2757  
Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTropLysGlyLeuLeuLysAspThrGlnThr 932  
Db 2758 GTTCAAGTGGCGGCCCAAGCTATTCCTGCTGGCGCTGTGCTGTGTGATACCGGAGCC 2817  
Qy 933 LeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2818 CTGAGGTGCAGAGCGCATCTACTGACATGTGCCGAGCTTCATCAGAGCCAGTCTCAC 2877  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2878 TTCAACCGCGCTTCAAGGTGGAGGAACATGCGCTCGAAACTTTTGGGGTCTTGGCG 2937  
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2938 CTGAAGTGCACAGCTGTGTTCTGGAATTTGCAAGTGAACAGCTTCACAGCGTGTGAC 2997  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 2998 AACATCTAACAAATCTCTGCTGACAGCGGTACAGTTTCAAGCATGTGTGTGACAGCTC 3057  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleGlnSerSerGln 1032  
Db 3058 CCATTTATCAGAAAGTTGGAAGAACCCCAACATTTTCTCGGCTCAATCTTACACG 3117  
Qy 1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3118 GCGTCCCTGTACTACTCATCTGAAAGCCAAAGACAGGAGATGCGTGGGGGCAAG 3177  
Qy 1053 GlySer-----PheProProGluAlaHisIleThrPheCysTyrGlnAlaPheLeu 1069  
Db 3178 GGGCGCCGCGCCCTCTGCTCCCTCCAGGCGCTGCAAGTGTGTGACACCAAGATCTCTG 3237  
Qy 1070 LeuLysLeuAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3238 CTCAAGTGACTGACACACGCTGACCTTACGAGCACTCTGGGTCACTCAGAGACAGC 3297  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3298 CAGACGCACTGTGTGTGAAGCTCCCGGGAGAGAGAGCTGACCTGAGAGCGCAGCC 3357  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3358 AACCCGCACTGCGCTCAGACTTCAAGACCAATCTGTGAC 3396

RESULT 5  
 BD094749  
 LOCUS  
 DEFINITION  
 The cell having the potentiality of differentiation into  
 cardiomyocytes.  
 ACCESSION  
 BD094749  
 VERSION  
 BD094749.1 GI:22640337  
 KEYWORDS  
 MO 0148150-A/16.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 3396)  
 Umezawa,A., Hata,J., Fukuda,K., Ogawa,S., Sakurada,K., Gojo,S. and  
 Yamada,Y.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 The cell having the potentiality of differentiation into  
 Patent: WO 0148150-A 16 05-JUL-2001,  
 KYOWA HAKKO KOGYO CO LTD,AKIHIRO UMEZAWA, JUNICHI HATA, KEIICHI  
 FUKUDA, SATOSHI OGAWA, KAZUHIRO SAKURADA, SATOSHI GOJO, YOJI YAMADA  
 OS Homo sapiens (human)  
 PN WO 0148150-A/16  
 PD 05-JUL-2001  
 PF 02-NOV-2000 MO 2000JP007741  
 PR 28-DEC-1999 JP 99P 372826, 28-FEB-2000 MO PCTJP0001448 PI  
 AKIHIRO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, PI  
 KAZUHIRO SAKURADA, SATOSHI GOJO, YOJI YAMADA  
 PC C12N5/06, C12N5/10, C12N15/09, A61K31/203, A61K35/28, A61K38/19, PC  
 A61K38/39,  
 PC A61K38/45, A61K48/00, A61P19/10, A61P41/00, C07K16/28, C12P21/08, PC  
 C12Q1/02,  
 PC C12Q1/48, G01N33/577  
 CC The cell having the potentiality of differentiation into CC  
 cardiomyocytes  
 FH Key Location/Qualifiers  
 FT CDS (1)..(3399).  
 Location/Qualifiers  
 1..3396  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 549 a 1157 c 1088 g 602 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,62e-174 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: Gaps: 13  
 US-09-042-460-2 (1-1122) x BD094749 (1-3396)  
 QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuValArgSerArgTyrArgGlu 20  
 DB 1 ATGGCGGCGGCTCCCGCTCGCGCGCGCTCGCTCGCTCGCGCGCGCGCGCGCGCGCG 60  
 QY 21 ValTTrpProLeuAlaThrPheValArgSerLeuGluProGluGluYArgArgLeuValGln 40  
 DB 61 GTGCTGCGGCTGCGCGCGCTGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCTGCTGCG 120  
 QY 41 ProGluAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 DB 121 CGCGGGGACCGCGGCGCTTCCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 180  
 QY 61 GlySerGlnProProPheAlaAspLeuSerPheHisGlnValSerSerLeuValGluLeu 80  
 DB 181 GAGCGACGGCG 240  
 QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 DB 241 GTGGCCCGAGTGTGCGAGAGGCTGTGCGAGCGCGCGCGCGAGAACTGTGCTGCGCG 300

QY 101 PheGluLeuLeuAsnGluValArgGlyCysProPheMetAlaPheThrSerSerValArg 120  
 DB 301 TTGGCGGTGTGCGAGCGGGCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 360  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValTTrpMetLeuLeu 140  
 DB 361 AGTAACTGTGCGCAACAGAGTGAACGCGACTGGGGGGGGGGGGGGGGGGGGGGGGGG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 DB 421 CTGGCG 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 DB 481 CTGGTGGCTCCACCTGCGCTGACAGTGTGGGGCGCGCGCGCGCGCGCGCGCGCGCG 540  
 QY 181 ThrThrAspIleTTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 DB 541 GCCACTACAGCG 594  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 DB 595 -----TGCGAACGGGCGCTGGAACCATGCTCAGGAGCGCGG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 DB 634 GTCCCGCTGGCGCTGCGAGCGCGGGTGCAGAGAGGGGGGGCGAGTGCACCGCAAGT 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 DB 694 CTGGCGGTGCCAAGAGGCGCGAGCGTGGCGCTGCGCTGACCGCGAGCGAGCGCGCTT 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 DB 754 GGGCAGGGGTCTGGCG 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
 DB 814 GTGGTGTACCTGCGCAG-----CCGCGCAGAGAACCACTCTTGGAG 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 859 GGTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 919 CCCCATTCACTGCG 978  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 979 GAGACCAAGCACTTCTCTACTCTCTCAGCGCAGC---AAGAGCAGCTGCGCGCGCTCTTC 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleLeuPhe 370  
 DB 1036 CTACTGAGCTCTTAGCG 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisAspLeuSerArgArg 390  
 DB 1096 CTGGGTTCAGCGCGCTGATGCCAGGAGCTCCCGCGAGTTGCCCGCGCGCGCGCGCG 1155  
 QY 391 TyrTTrpGlnMetArgProLeuPheGlnIleLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1156 TACTGGCAATGCG 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 DB 1216 GGGGTCTCTCAAGACGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1266  
 QY 431 Leu-----Asn 452  
 DB 1267 GCCGATGTGTGCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 1326  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTTrpGlnValTyr 452

Db 1327 ACAGACCCCGCTCCGCTGCTGACAGCTGCTCCGCGACAGACAGCCCTCGGACAGGTTAC 1386  
 Qy 453 G1yPheLeuAArgAlaCyLeuCyLeuValValSerAlaSerLeuTPrGlyThrArgHis 472  
 Db 1387 GGGCTTCGTGGGGCTCCCTCCGCGCGCTGCTGCTCCGCGCGCTCCGCGCGCTCCGCGCGC 1446  
 Qy 473 AAmG1uAArgAArgPhePheLeuValValSerPheLeuLeuLeuLeuLeuLeuLeuLeuLeu 492  
 Db 1447 AACGACGCGCGCTCCCTCCGCGCGCGCTGCTGCTCCGCGCGCTCCGCGCGCTCCGCGCGC 1506  
 Qy 493 LeuSerLeuG1uLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 512  
 Db 1507 CTCTCGCTGCGAGAGCTGACGCTGAGATAGCGCTGCGAGAGCTGCGCGCTGCGCGCGCAG 1566  
 Qy 513 SerProG1yLysAArgValProAlaAlaG1uHisAArgPheArgG1uArgG1uLeuAla 532  
 Db 1567 AGCCGAGGGGTTGGCTGTGTTCCGCGCGAGACCCGCTGCGTGAAGAGATCTGAGCC 1626  
 Qy 533 ThrPheLeuPheTPrLeuMetAArgPheTyrValValG1uLeuLeuArgSerPhePheTyr 552  
 Db 1627 AAGTTCTGCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686  
 Qy 553 IleThrG1uSerThrPheG1uLeuValValValValValValValValValValValValVal 572  
 Db 1687 GTCAAGGAGACAGAGGTTTCAAAAGAGAGCTCTTTTCTACCGGAGAGGTCTCTGAGAC 1746  
 Qy 573 LysLeuG1uSerIleG1yValArgG1uHisIleuG1uArgValArgLeuArgG1uLeuSer 592  
 Db 1747 AAGTTGCAAAAGCATTTGAATAGACACAGCATTTGAAGAGGTGAGCTGCGGAGAGCTGTG 1806  
 Qy 593 G1uG1uG1uValAArgHisIleG1uSerThrTPrLeuAArgProIleCyAArgLeuArg 612  
 Db 1807 GAAGGAGAGGTCAAGGAGCATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1866  
 Qy 613 PheIleProLysProAArgG1yLeuArgProIleValAArgMetSerTyrSerMetG1yThr 632  
 Db 1867 TTCAATCCCAAGCTCGAGGAGCTGCGGCGCATTTGTAACATGACATGCGCGGAGAGAGC 1926  
 Qy 633 ArgAlaLeuG1yAArgArgG1yG1uAlaG1uHisPheThrG1uArgLeuLysThrLeuPhe 652  
 Db 1927 AGAAGGTTCCGAGAGAAAG 1986  
 Qy 653 SerMetLeuAArgG1uArgThrLysHisPheHisPheHisPheHisPheHisPheHisPheHis 672  
 Db 1987 AGCGTCTCACTAG 2046  
 Qy 673 MetAAspIleTyrArgThrTPrAArgAlaPheValLeuArgValArgAlaLeuAAspGln 692  
 Db 2047 CTGACAGATATCCAGAGGAGCTGCGGAGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106  
 Qy 693 ThrProAArgMetTyrPheValValValValValValValValValValValValValVal 712  
 Db 2107 CCGCTGAGCTGATCTTGTCTCAAGGTGATGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2166  
 Qy 713 G1yLysLeuValG1uValValAlaAAspMetIleAArgHisSerG1uSerThrTyrCyAile 732  
 Db 2167 GACAGGCTCAGGAGGTGATGCGGAGCATCAAA---CCCAAGAACAGTACTGCGTG 2223  
 Qy 733 ArgG1uTyrAlaValValAArgAArgAAspSerG1uG1uValHisIleLysSerPheAArg 752  
 Db 2224 CGTCGATATGCGGTGCTCAG 2283  
 Qy 753 G1uValThrThrLeuSerAAspLeuG1uProTyrMetG1yG1uPheLeuLysHisIleG1u 772  
 Db 2284 CAGCTCTCACTTACAGACCTCAGGCGGTAATGACAGAGTTCGAGGCTCAGCTCAGAG 2343  
 Qy 773 AAspSerAAspAlaSerAlaLeuArgAAspSerValValIleG1uG1uSerIleSerMetAAsn 792  
 Db 2344 GAGAGC---AGCCCGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2397  
 Qy 793 G1uSerSerSerSerLeuPheAAspPhePheLeuHisPheLeuAArgHisSerValValLys 812

Db 2398 GAGCGCAGAGTGGCTCTTCGACGCTTCTACGCTTCATGTCACACAGCCGTTGCGC 2457  
 Qy 813 IleG1yAAspArgCySerThrThrG1uG1uG1uG1uIleProG1uG1uSerSerLeuSerThr 832  
 Db 2458 ATCAAGGGGAGAGCTTACGCTCAG 2517  
 Qy 833 LeuLeuCySerLeuCyAAspMetG1uAAspLeuValValValValValValValValValVal 852  
 Db 2518 CTGCTGCAAGCTGTGCTGAGGAGCATGAGAGAACAGCTGTTGCGGAGAGAGAGAGAGAGAG 2577  
 Qy 853 AAspG1yLeuLeuLeuArgPheValAAspAArgPheLeuValThrProHisIleAAspGln 872  
 Db 2578 GACGGGTGCTCTGCGTTGTTGGATGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2637  
 Qy 873 AlaLysThrPheLeuSerThrLeuValHisG1yValProG1uTyrG1yCyAArgIleAAsn 892  
 Db 2638 GCGAAACCTTCTCAG 2697  
 Qy 893 LeuG1uLysThrValValAAspPheProValG1uProG1yThrLeuG1yG1yAlaAlaPro 912  
 Db 2698 TTGCGGAGACAGTGTGTAATTCCTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2757  
 Qy 913 TyrG1uLeuProAlaHisCySerLeuPheProTPrCyG1yLeuLeuLeuAAspThrG1uThr 932  
 Db 2758 GTTCAGATGCGGAG 2817  
 Qy 933 LeuG1uValPheCyAAspTyrSerG1yTyrAlaG1uThrSerIleLysThrSerLeuThr 952  
 Db 2818 CTGAGAGGTGAG 2877  
 Qy 953 PheG1uSerValPheLysAlaG1yLysThrMetAArgAAspLysLeuLeuSerValLeuArg 972  
 Db 2878 TTCAACCGCGCTTCAAGGTGAG 2937  
 Qy 973 LeuLysCyHisIleG1yLeuPheLeuAAspLeuG1uValAAspSerLeuG1uThrValCyAile 992  
 Db 2938 CTGAAGGTGACAGAGCTGTTCTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2997  
 Qy 993 AAspIleTyrIlePheLeuLeuG1uAlaTyrAArgPheHisAlaCyAValIleG1uLeu 1012  
 Db 2998 AAcATTCAGAGAGCTCTGCTGAG 3057  
 Qy 1013 ProPheAAspG1uAArgValAArgLysAAspLeuThrPheLeuG1yIleIleSerSerGln 1032  
 Db 3058 CCAATTCATCAGAGAGAGTGGAG 3117  
 Qy 1033 AlaSerCySerTyrAlaIleLeuLysValValAAspProG1yMetThrLeuLysAlaSer 1052  
 Db 3118 GCTTCCCTCTGCTACTCCATCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3177  
 Qy 1053 G1ySer-----PheProProG1uAlaAlaHisIleTPrLeuCySerTyrG1uAlaPheLeu 1069  
 Db 3178 GGGCGCGCGGCTCTGCGCTCCGAGAGCGGTCAGATGCTGTGCGACACAGATTCCTCG 3237  
 Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCyLeuLeuG1yProLeuAArgThrAla 1089  
 Db 3238 CTCAAGCTGACTCAGACACCGGTGACCTTACGTGCACTCTGCGGAGTCACTCAGAGAGAGC 3297  
 Qy 1090 G1uLysLeuLeuCyAArgLysLeuProG1uAlaThrMetThrIleLeuLysAlaAlaAla 1109  
 Db 3298 CAGAGCAGCTGAGTCCAGC 3357  
 Qy 1110 AAspProAlaLeuSerThrAAspPheG1uThrIleLeuAAsp 1122  
 Db 3358 AACCGGAGACTGCCCTCAGACTTCAAGAGACATCTCTGAGC 3396  
 RESULT 6  
 BD096291 3396 bp DNA linear part 27-AUG-2002  
 LOCUS  
 DERIVATION Cells capable of differentiating into myocardial cells.  
 ACCESSION BD096291  
 VERSION BD096291.1 GI:22641879  
 KEYWORDS MO 0148151-A/16.



SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 3396)  
 AUTHORS Umezawa, A., Hata, J., Fukuda, K., Ogawa, S., Sakurada, K., Gojo, S. and  
 Yamada, Y.  
 TITLE Cells capable of differentiating into myocardial cells  
 JOURNAL Patent: WO 0148151-A 16 05-JUL-2001;  
 KYOWA HAKKO KOGYO CO LTD  
 COMMENT OS Homo sapiens (human)  
 PN WO 0148151-A/16  
 PD 05-JUL-2001  
 PR 27-DEC-2000 WO 2000JP009323  
 BR 28-DEC-1999 JP 99P 372826, 28-FEB-2000 WO PCTJP0001148 PR  
 02-NOV-2000 WO PCTJP0007741  
 PI AKIHIRO UMEZAWA, JUNICHI HATA, KEIICHI FUKUDA, SATOSHI OGAWA, PI  
 KAZUHIRO SAKURADA, SATOSHI GOJO, YOTI YAMADA  
 PC  
 C12N5/06, C12N5/08, C12P21/08, C12Q1/02, A61K35/28, A61K35/44, A61P9/ PC  
 06', A61P9/04//A61K38/18, C12N5/12  
 CC Cells capable of differentiating into myocardial cells FH  
 Key Location/Qualifiers  
 FT CDS Location/Qualifiers  
 1. 3396  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 BASE COUNT 549 a 1157 c 1088 g 602 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,62e-174 Length: 3396  
 Score: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 DB: Gaps: 13  
 US-09-042-460-2 (1-1122) x BD096291 (1-3396)  
 QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20  
 Db 1 ATGCCGGCGCTCCCGCTGCCAGCGCTGCCCTGCTGCCAGCCACTACCGCGAG 60  
 QY 21 ValTrpProLeuAlaThrPheValaArgLeuGlyProGluGlyArgArgLeuValGln 40  
 Db 61 GTGCTGCCGCTGCCACGCTTCGTCGGCGCGCTGCCGCCAGGGCTGGCGCTGGTCAG 120  
 QY 41 ProGlyAspProIysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 Db 121 CGCGGGAGCCCGCGGCTTCCGCGCTGGTGGCCAGTGCCTGGTGGCTGGCTGG 180  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuArgGluLeu 80  
 Db 181 GAGCAGCGCGCGCGCGCGCGCGCGCGCTTCCGCGAGTGTCTGCTGAAGAGCTG 240  
 QY 81 ValAlaArgValAlaGlnArgLeuCysGlyArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 241 GTGCCCGGAGTGTGCAGAGCTGTGTGCAGCGCGCGCGAGAACCTGTGCTGCCTCGGC 300  
 QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 Db 301 TTCGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGAGGCTTCCACACGAGCTGGCC 360  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 Db 361 AGTACCTGCGCAACGAGTGCAGACGCTGCGGGGAGCGGGGCGCTGGCGCTG 420  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysValaLeuTyrLeu 160

Db 421 CTGGCGCGCTGGCGAGCAGCGTGTGTTCACTGCTGAGCAGCTGCGCGCTTGTG 480  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 481 CTGGGTCTCCACCTGGCTGCTACAGAGTGTGGGCGCGCTGTACACAGCTCGCGCT 540  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 541 GCCACTGAGCGCGCGCGCGCGCGCGCGAGCTAGT---GACCCCGAAGCGCGTGGGA--- 594  
 QY 201 AsnPheThrAsnLeuAlaPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 Db 595 -----TGCAGAGCGCGCTGGAACCATGCTGCAGAGAGCGCGG 633  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 634 GTCCCCCTGGGCGCTGCCAGCGCGCGGTCGAGAGAGGCGCGGCGAGTGCAGCGAGT 693  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 694 CTGCCGTTGCCCAAGAGCGCGCGCGCTGGCGCTGCCCTGAGCGGAGCGGACGCGCTT 753  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 754 GCGCAGGAGTCTGAGCGCCACCGCGGAGAGCGGTGACCGAGTACCGTGGTTCTGT 813  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 814 GTGGGTCACTGCAGCA-----CCCCCGAAGAGCACCTTTGGAG 858  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 859 GGTGGCTCTCTGGCAGCGCGCATCTCCACCATCTGAGCGCGCGCGAGCACCGCGGCG 918  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 919 CCCCACCATCAGTCGCGGCGCACAGTCCCTGGAGACAGCGCTTCCCGCGTGTAGCC 978  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 979 GAGACCAAGCACTTCTTACTTCTCTCAGGACAC---AAGAGCAGCTCGGCTCTCTTC 1035  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370  
 Db 1036 CTACTCAGTCTCTGAGCCGACAGCTGAGCTGCGCTGAGAGCTCGTGAGACCATCTTT 1095  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuLysArgThrHisArgLeuSerArgArg 390  
 Db 1096 CTGGGTTCAGGCGCTGATGCCAGGAGCTCCGCGAGGTGGCCCGCGCTGCCAGCGC 1155  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1156 TACTGGCAATCGCGCGCTGTTCTGTGAGTGTGTTGGAAACAGCGGAGTGGCCCTTAC 1215  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1216 GGGGTCTCTCAAGACCACTGCGCGCTGCGAGCTGG---GTCAACCCAGCA 1266  
 QY 431 Leu-----Asn 432  
 Db 1267 GCCGCTGTCTGCGCGGAGAACGCCCAAGGCTCTGCGCGCGCGCGCGAGAGAGAGAC 1326  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1327 ACAAGACCCCGTGGCTGTGAGCTGTCCGCGACAGCACACACCCCTGCGAGTGTAC 1386  
 QY 453 GlyPheLeuArgAlaCysLeuLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1387 GGGTGTGGGCGCTGCTGCGCGCTGAGTGTGCGCGCGCGCGCTTGGGCTCCAGGAC 1446  
 QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
 Db 1447 AACGAACGCGCTTCTCTCAGAAACACCAAGATTCACTCTCGGGGAGAGCATGCCAAG 1506



```

source      1..3399
organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
SE COUNT    550 a 1157 c 1089 g 603 t
IGIN

ignment Scores:
ad. No.:      3,628-174      Length:      3399
pre:          3505.00      Matches:      719
cent Similarity: 72.94%      Conservative: 122
st Local Similarity: 62.36%      Mismatches: 260
ary Match:      59.40%      Indels:      52
:               6          Gaps:      13

-09-042-460-2 (1-1122) x AX481414 (1-3399)
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
1 ATGCCGGCGCTCCCGCTCCGAGCGCTGGCTCCCTGCTGGCGAGCCACTACCGCGAG 60
21 ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValcIn 40
61 GTGCTCCGCTGGCCACGTTCGTGGCGGCTGGGGCCCGCCAGGCTGGCGGCTGGTGCAG 120
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
121 CGCGGGAGACCCGGCGCTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGTGGCTGG 180
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyGluLeu 80
181 GAGCGACGGCGCGCCCGCCCGCTCCCTCCGCGAGGTGCTCGCTGAGGAGGTG 240
81 ValAlaArgValValGlnArgLeuGlyGluArgAsnGluArgAsnValLeuAlaPheGly 100
241 GTGCGCGAGTGTGAGAGAGCTGTCCGAGCGCGCTGGCGGAGGAGAGAGTGTGGCTTCGGC 300
101 PheGluLeuLeuAlaAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
301 TTCGCGCTGTGGACGGCGCGCGGGCGCCCGCCCGAGGCGCTTCACCAACAGGCGTGGCG 360
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
361 AGTACTGTGCCAACACGGTGCACGACGACCTGCGGGGAGCGGGGCGTGGGGGTGGTGT 420
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
421 CTGCGCGCGCTGGCGAGCGAGCTGTGCTTACCTGTGCGACGCTGCGCGCTCTTTGTG 480
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
481 CTGGTGGCTCCACAGCTCGCCTACCAAGTGTGGCGGCGCGCGCTGTACAGCTGGCGCT 540
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
541 GCACACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
595 -----TGCAGACGGCGCTGGAACCATAGCGTCCAGGAGCGCGGG 633
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
634 GTCCCTCTGGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 693
241 ValProSerAlaLysAlaLysCysTyrProValProArgValGluGlyPro--- 259
694 CTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 753
260 -----HisArgGlnValLeuProThrProThrProSerGlyLysSerTyr--- 272
754 GGGCAGGGGTCTCTGGGCGCCACCGCGGAGGACGCGTGGAGCCGAGTACCGTGGTTCGTGT 813

```

```

QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
DB 814 GTGTGTCACTGCTCCAGA-----CCGCCGAGAGAGCCACCTCTTTGGAG 858
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
DB 859 GGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGCGCGCCAGCACCACCGCGGC 918
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 919 CCCCCATCCACATCGCGGCCACACGCTCCCTGGGACAGCCCTTGTCCCGGTGTACGCC 978
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
DB 979 GAGACCAAGCACTTCTCTACTCTCTCAGCGGAC---AAGGAGCAGCTGGCGGCTCTCTTC 1035
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
DB 1036 CTACTCAGCTCTCTGAGCGCCAGCGCTGACTGGCGCTGGAGGCTCTGGAGACCATCTTT 1095
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
DB 1096 CTGGGTTCCAGCGCTCCGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCGC 1155
QY 391 TyrTyrGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
DB 1156 TACTGGCAATGCGCGCCCTGTTTCTGGAGCTGTTGGGAACACCGCGCAGTGCCTCTAC 1215
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
DB 1216 GGGGTGCTCTCAAGACGACTGCCGCTGGAGCTGCG-----GTACCCCGCAGCA 1266
QY 431 Leu-----Asn 432
DB 1267 GCGGTGTCTGTGCCGGGAGAGCCCGCGGCTCTGTGGCGGCCCGCGAGGAGGAG 1326
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452
DB 1327 ACAGACCCCGCTGCTGGTGTGCTGCTGCCGAGCAGCAGCAGCCCTGGCAGGTGTAC 1386
QY 453 GlyPheLeuAlaGlyAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472
DB 1387 GACTTCTGGTGGGCGCTGCTGCTGCCGCGGTGTGCTGCCCGCGCTCTGGGCTCCAGGCAC 1446
QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492
DB 1447 AACGAAACGCGCTTCTCAGGACACCAAGAAAGTTCATCTCCCTGGGGAAGCATGCCAAG 1506
QY 493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512
DB 1507 CTCTGCTGCAGGAGCTGACGTGGAGATGAGCTGCGGAGCTGCGCTTGGCTGGCGCAGG 1566
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlyLeuAla 532
DB 1567 AGCCAGCGGTTGGCTGTGCTTCCGCGCGCAGAGACCGCTCTGCGTGGAGGATCTCGGCC 1626
QY 533 ThrPheLeuPheTyrProLeuMetAspThrTyrValGlnLeuLeuArgSerPhePheTyr 552
DB 1627 AAGTTCCTGCACTGGCTGATGAGTGTGCTGCTGCTGAGCTGCTCAGGTCTTCTTTAT 1686
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572
DB 1687 GTCAGGAGACCACTTTCAAAGAAAGAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1746
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
DB 1747 AAGTTGCAAGCATTTGGAATCAGACAGCATGAGAGGGTGGAGCTGCGGAGACTGTGCG 1806
QY 593 GlnGluGlnValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgLeuArg 612
DB 1807 GAAGCAGAGGTTCAGGACGATCGGAGAGCCAGGCGCGCTGCTGCTGCTGCTGCTGCTG 1866
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632

```

1867 TTAATCCCAAGCTGACGGCTGCGGCCGATTTGTAACATGAGCTACCTGCTGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAAGTTCCGACAGAAAGAGGGCCGCGCTCACCTCGAGGGTGAAGCACTGTTC 1986  
553 SerMetLeuSerThrGluArgThrLysHisPheProHisLeuMetGlySerSerValLeuGly 672  
1987 AGCGTGTCAACTACGACGGCGCGCGCCCGCGCTCTCTGGCGCGCTCTGTGTGGGC 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGACAGATATCCACAGGGCTGGCGACCTTCGTGCTGGTGGCGGCCCGACAGACCG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CCGCTCGAGCTGACTTGTCAAGGTGATGTGACGGCGCGCTACGACACCATCCGCCAG 2166  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACAGGCTCAGGAGTTCATCGCCAGCATCATCAA---CCCGAACACAGTACTGGTG 2223  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 CGTCGGTATGCGGTGTCAGAAAGCGGCCCATGGCGACGTCGCGCAAGCCCTCAAGAGC 2283  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CACGTCTCTACCTTGACAGACCTCCAGCGCTCATCGCAGTTCGTGGCTCACCTGACG 2343  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2344 GAGACC-----AGCCCGCTGAGGATGCGCTGTCATCGACGAGAGCTCCTCCCTGAT 2397  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAGCCACAGAGTGGCTCTTCGACGCTCTCTACGCTTCATGTGCCACACCGCGTGGCG 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCAGGGCAGTCTTACGTCCAGTCAGTCAGCGGATCCCGAGGGCTTCATCTCTCCAG 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 CTGCTCTGACGCTGTGCTACGGCGCATGAGAAAGTGTTCGGGGATTCGGCGG 2577  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872  
2578 GACGGCTGCTCCTCGCTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2637  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 CGCAAAACCTTCTCCAGACCTGTGTCGAGGTGTCCTGAGTATGCTGCTGGTGTGAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912  
2698 TTGGGAAGACAGTGGTGAATCTCCCTGTAGAAGACGAGGCCCTGGGTGGCGGCTTT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATGCGGCCCGCCCGGCTATCTCCCTGGTGGCGCTGTGTGTGATACCCGGACC 2817  
933 LeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGGAGGTGCAGAGCGACTATCCAGCTATGCGCGACCTCCATCAGACCACTCTCAC 2877  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGCGCTTCAAGGCTGGAGGAACATGCTGCGCAAACTCTTTGGGTCTTGGCG 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992

Db 2938 CTGAAGTGTCAAGCCTGTTTCTGATTGCGAGTGAACAGCTCCACAGCGTGTGCACC 2997  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 2998 AACATCTACAGATCTCTGCTGCGAGCGTACAGGTTTCACGATGTGTGTCAGCTC 3057  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3058 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTGCGGCTCATCTCTGACACG 3117  
Qy 1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3118 GCCTCCTCTGCTACTCCATCTCTGAAGCAAGACGAGGATGTCGCTGGGGGCCAAG 3177  
Qy 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
Db 3178 GCGCGCGCGCGGCTCTGCGCTCCGAGGCGGTGAGTGGCTGTGCCACCAAGCATTCCTG 3237  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3238 CTCAGCTGACTCGACACCGTGTCACTACGTGCCACTCTGGGTCACTCAGGACAGCC 3297  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
Db 3298 CAGACGAGCTGAGTCGGAAGCTCCCGGGAGACGCTGACTGCCCTGGAGCGCGACCC 3357  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3358 AACCCGGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3396  
RESULT 8  
ARI04587  
LOCUS ARI04587 4015 bp DNA linear PAT 14-FEB-2001  
DEFINITION Sequence 224 from patent US 6093809.  
ACCESSION ARI04587  
VERSION ARI04587.1 GI:12817295  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Cecch,T.R. and Lingner,J.  
TITLE Telomerase  
JOURNAL Patent: US 6093809-A 224 25-JUL-2000;  
FEATURES Location/Qualifiers  
source 1. .4015  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,54e-174 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 6 Gaps: 13  
US-09-042-460-2 (1-1122) x ARI04587 (1-4015)  
Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGGCTCCCGCTCCCGCGCGCGCTGCTGCTGCTGCGAGCCACTACCGGAG 115  
Qy 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGGCCACGTTCTGCGGGCGCTGGGGCGCCCGAGGCTGGCGGTGGTGCAG 175  
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGGACCCCGCGGCTTTTCCGCGCGCTGGTGGGCCAGTGCCTGTGTGCTGCTTGG 235  
Qy 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

```
236 GAGCAGCGCCGCCCGCCGCCGCCCTCTTCCGCCAGGCTGCTCCCTGCAAGGAGCTG 295
      :::::|||||
81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
      :::::|||||
296 GTGGCCCGAGTCTGAGAGGCTGTGCGAGCGCGCGGCGAAGACGTGCTGGCCTTCGGC 355
      :::::|||||
101 PheGluLeuLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
      :::::|||||
356 TTGGCGCTCTGGACGGGCGCGCGGGCGCCGCCCGAGGCTTTCACCAACAGCGTGGCG 415
      :::::|||||
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
      :::::|||||
416 AGCTACTGCCCAACAGGTGACCGACGCTGCGGGGACGCGGGGTGGGGCTGCTG 475
      :::::|||||
141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
      :::::|||||
476 CTGGCGCGGTGGGCGACGACGTGCTGGTTACCTGCTGGCAGCGTGGCGCTCTTTGTG 535
      :::::|||||
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
      :::::|||||
536 CTGGTGGCTCCAGCTGCCCTTACAGGTGTGCGGCGCGCGCTGTACCAAGCTCGGCGCT 595
      :::::|||||
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
      :::::|||||
596 GCCACTCAGCGCGCGCCCGCCACACGCTAGT---GGACCCCGAAGCGCTCGGA--- 649
      :::::|||||
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
      :::::|||||
650 -----TGCAGACGGGCTGGAACTTACCTAGCTCAGGAGGCGCGG 688
      :::::|||||
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240
      :::::|||||
689 GTCCCTCTGGGCTGCGACCCCGGCTGCGAGGAGCGCGGGCGAGTGCACCGAAGT 748
      :::::|||||
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
      :::::|||||
749 CTGCCCTTCCCAAGAGGCGCGGCGTGGCGTGCCTCGAGCGGAGCGGACCCCGTT 808
      :::::|||||
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
      :::::|||||
809 GGCAGGGGTCTGGGCCACCGCGGCGAGCGCGTGGACCGAGTACCGGTGTTCTGT 868
      :::::|||||
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
      :::::|||||
869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGACCCACCTCTTTGGAG 913
      :::::|||||
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
      :::::|||||
914 GGTGGCTCTCTGGACGGCGCACTCCACCATCCGTGGCGCGCCACGACACCGCGGCG 973
      :::::|||||
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
      :::::|||||
974 CCCCCATCCACATCGCGGCGCACACGCTCCCTGGGACAGCGCTGTGCCCGGTAGCGC 1033
      :::::|||||
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
      :::::|||||
1034 GAGCAAGCAAGCACTTCTCTACTCTCAGCGGAC---AAGGAGCACTCGGCGCTCTTC 1090
      :::::|||||
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370
      :::::|||||
1091 CTACTCAGCTCTCTAGGCCCGACCTGACTGCGGCTCGGAGGCTCGTGAGACCATCTTT 1150
      :::::|||||
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
      :::::|||||
1151 CTGGGTCCAGGCCCTGGATGCGCAGGACTCCCCCGAGGTGCCCCGCTGCCCGCGCGC 1210
      :::::|||||
391 TyrTriPLeuMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
      :::::|||||
1211 TACTGGCAATGCGGCCCTGTTCCTGGAGCTGCTTGGGAACCAACGCGAGTGGCCCTAC 1270
      :::::|||||
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
      :::::|||||
```

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerMetAsn 792  
: : : : :  
2399 GAGACC-----AGCCCGCTGAGGATGCGGTGTCATCGAGCAGAGCTCTCCCTGGAAT 2452  
: : : : :  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
: : : : :  
2453 GAGGCCAGAGTGGCCCTTCGACGTCTTCTACGCTTCATGTGCCACACGCGGTGGC 2512  
: : : : :  
813 IleGlyAspArgCysTyThrGlnCysGlnGlyLeProGlnGlySerSerLeuSerThr 832  
: : : : :  
2513 ATCAGGGCAAGTCTCATGTCAGTCAGGAGATCCCGCAGGCTCCATCTCTCCAGC 2572  
: : : : :  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
: : : : :  
2573 CTGCTCTGACACCTGTGTCGCGGACATGAGAACAGCTGTTCGGGGATTCGGCGG 2632  
: : : : :  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
: : : : :  
2633 GACGGCTGCTCTGCGTTTGGTGGATGATTTCTTTGTGTGACACCTCACCTCACCCAC 2692  
: : : : :  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyArgCysMetIleAsn 892  
: : : : :  
2693 GCGAAACCTTCTCAGACACCTGTCGAGGTGTCCCTCAGTATGCTGCTGATACCCGACC 2752  
: : : : :  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
: : : : :  
2753 TTGCGGAAGACAGTGGTGAATCTTCTCTGTAGAACGAGGCGCTGGGTGGCAGCGCTTT 2812  
: : : : :  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
: : : : :  
2813 GTTCAGATGCGCGGCCACCGGCTATTCCCTGTGGTGGCGCTGCTGTGATACCCGACC 2872  
: : : : :  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
: : : : :  
2873 CTGAGGTGCGAGCGACTACTCAGCTATGCGCGACCTTCATCAGGCCAGTCTACCC 2932  
: : : : :  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
: : : : :  
2933 TTCAACCGCGCTTCAAGGCTGGGAGGAACATGCTGCGCAATCTTTGGGTCTTGGCG 2992  
: : : : :  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGluValAsnSerLeuGlnThrValCysIle 992  
: : : : :  
2993 CTGAAGTGTACAGCGTGTCTCGATTTGAGGTGAACGCTCCAGACGGTGTGCACC 3052  
: : : : :  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
: : : : :  
3053 AACATCTACAAGATCTCTGCTGTCAGCGGTACAGGTTTCAACGATGTGTCTGCAGCTC 3112  
: : : : :  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyLysIleSerSerGln 1032  
: : : : :  
3113 CCATTTCTATCAGCAAGTTTGAAGAACCCACACATTTTCTGCGCGTCTATCTGTACACG 3172  
: : : : :  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
: : : : :  
3173 GCCTCCCTCTCTACTCTCATCTCTGAAGACCAAGACCCAGGAGTGTGCTGGGGCCAG 3232  
: : : : :  
1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
: : : : :  
3233 GCGCGCGCGCGCCCTCTGCGCTCCGAGCGCGCTGAGTGGTGTGCCACCAAGCATTTCTG 3292  
: : : : :  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
: : : : :  
3293 CTCAGCTGACTCAGACCGGTGTCACTAGTCCACTCTCTGGGGTCACTCAGGACAGCC 3352  
: : : : :  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
: : : : :  
3353 CAGACCGAGCTGAGTCGGAAGTCTCCCGGGACGACGCTGACTCTCCCTGTGAGCGCGAC 3412  
: : : : :  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
: : : : :  
3413 AACCCGGCAGCTGCGCTCAGACTTCAAGACCATCTCTGGAC 3451

## RESULT 9

ARI175848  
LOCUS ARI175848 4015 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 224 from patent US 6309867.  
ACCESSION ARI175848  
VERSION ARI175848.1 GI:17917147  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Cech,T.R. and Nakamura,T.  
TITLE Telomerase  
JOURNAL Patent: US 6309867-A 224 30-OCT-2001;  
FEATURES Location/Qualifiers  
1..4015  
/organism="unknown"  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,54e-174 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 6 Gaps: 13  
US-09-042-460-2 (1-1122) x ARI175848 (1-4015)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCCGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115  
QY 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGGTGCGCACGTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
Db 176 CGCGGGACCGCGCGCTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 235  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGlnLeu 80  
Db 236 GACGACGCGCGCGCGCGCGCGCGCTCTCTTCCGCCAGGTGCTCTGCTGCTGCTGCTG 295  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGCGCGAGTGTGCGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTG 355  
QY 101 PheGluLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
Db 356 TTGCGCGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTCAACACAGCTGCG 415  
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140  
Db 416 AGTACTGCTGCCAACCGGTGACCGACGACCTGCGCGCGCGCGCGCGCGCGCTGCTGCTG 475  
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
Db 476 CTGCGCGCGTGGCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
Db 536 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595  
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
Db 596 GCACCTCAGCTGCTGCTG 649  
QY 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerSerArgGlnGluAlaPro 220  
Db 650 -----TCCGAACCGCGCTGGAACCATAGCTCAGGCGCGCGCGCGCGCGCGCGCG 688

QY	553	IleThrGlnSerThrPheGlnLysAsnArgLeuPheThrArgLysSerValTrpSer	572
DB	1742	GTACGGAGACACAGCTTTCAAAGAACAGAGCTCTTTTCTACCGGAAGAGTGTCTGGAGC	1801
QY	573	LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer	592
DB	1802	AGTTGCAAGAGCATTTGGAATCAGACAGCACATTGAAGAGGGTGCAGCTCGGGAGGTGTGG	1861
QY	593	GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg	612
DB	1862	GAAGCAGAGGTCTCAGGCAGCATCGGAAGACAGCGCCCGCTGCTCAGCTCCAGACTCCGC	1921
QY	613	PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr	632
DB	1922	TTCTATCCCAAGCCTGACGGGCTGGCGCATTTGTGAACATGGACTACTCGTCGTGGAGGCC	1981
QY	633	ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe	652
DB	1982	AGAACGTTCGGCAGAGAAAGAGGGCCGAGCTCTCACCTCAGAGGGTGAAGCACTGTTC	2041
QY	653	SerMetLeuAsnTyrGlnArgThrLysHisProHisLeuMetGlySerSerValLeuGly	672
DB	2042	AGCGTGCTCAACTACAGCGGGCGGGCCCGGCTCTCTGGGCGCCTTCTGTGTGGGC	2101
QY	673	MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln	692
DB	2102	CTGAGCATATCCACAGGGCTGGCGCACCTTCGTGCTGCGTGTGGGGCCCGAGACCCG	2161
QY	693	ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln	712
DB	2162	CCCGCTGAGCTGTACTTTGTCAAGGTGGATGTACGGGGCGGTACACACCATCCCCAG	2221
QY	713	GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle	732
DB	2222	GACAGGCTCAGGAGGTCTATCGCCAGCATCATCAA---CCCCAGAACAGTACTGCGTG	2278
QY	733	ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg	752
DB	2279	CGTCGGTATCGCTGTGTGAGAAAGGCGGCCATGGCCACGTCCGCAAGGCCCTTCAAGAGC	2338
QY	753	GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln	772
DB	2339	CACGTCTCTACCTTGACAGACCTCCACCGCATCATCGACAGTTCGTGGCTCACCTGCAG	2398
QY	773	AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnGlnSerIleSerMetAsn	792
DB	2399	GAGACC-----AGCCCGCTGAGGATGCCGTGCTCATCGAGCAGAGCTCTCCCTGAAT	2452
QY	793	GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys	812
DB	2453	GAGCCACGACGTGGCTCTTCGACGCTTCCTCAGCTTCATGTGCCACACGCCGCTGGCC	2512
QY	813	IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerLeuSerThr	832
DB	2513	ATCAGGGGCAAGTCTTACGTCAGTGCCAGGGGATCCCGCAGGGCTCCATCTCTCCACG	2572
QY	833	LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg	852
DB	2573	CTGTCTGCACTGTCTGTCAGCGCATATGGAGAACAAAGCTGTTTCGGGGATTCGGCGG	2632
QY	853	AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln	872
DB	2633	GACGGGCTGCTCCCTGGTTGGTGGATGATTTCTTGTGTGTGACACCTCACCTCACCCAC	2692
QY	873	AlaLysThrPheLeuSerThrLeuValHisGlyValProGlnTyrGlyCysMetIleAsn	892
DB	2693	CGCAAAACCTTCCTCAGAGCCCTGGTCCGAGGTGTCTCCTGAGTATGCTGCTGGGTGAAC	2752
QY	893	LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro	912
DB	2753	TTGGGAGAGACAGTGGGTGAATCTTCCCTGTAGAGACGAGGCCCTTGGGTGCACGGCTTTT	2812
QY	913	TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr	932



```

2813 GTTCAGATCGCGCCACCGCTATTCCTCGTGGCGCTGCTGCTGATACCGGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGAGAGTGCAGAGCAGTACTCCAGCTATGCGCGAGCTTCCATCAGAGCCAGTCTCACC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972
2933 TTCAACCGCGGCTTCAAGCTGGAGAACATGCTCGCAACTCTTTGGGGTCTTCGGG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGACCTGTTCTGATTTGCAGGTGAACAGCGCTCCAGACGCTGTGCACC 3052
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAGATCCTCTGCTGAGCGGTACAGTTTCAGGATGTGCTCGAGCTC 3112
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTCGCGCTCATCTCTGACACG 3172
1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCTCCTCTGCTACTCTCATCTGAAAGCCAAAGACGAGGAGTGTGCTGGGGGCCAAG 3232
1053 GlySer-----PheProGlnAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGGCGCGCGCCCTGCTCGCTCGAGCGCGTGCAGTGTGCTGCCACCAAGCATCTCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAGCTGACTCGACACCGCTGCACCTACGTGCGACCTCTCGGGGTCACTCAGACAGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGlnAlaThrMetThrIleLeuLysAlaAla 1109
3353 CAGACCGACTGATCGGAAGCTCCGGGAGCAGCGCTGCTGCCCTGGAGGCGCAGCC 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCCGCGCTGCTCGCTCAGACTTCAAGACCATCTCTGGAC 3451

```

## SULT 10

```

CUS AR182221 4015 bp DNA linear PAT 20-APR-2002
FINITION Sequence 1 from patent US 6337200.
CESSION AR182221
RSION AR182221.1 GI:20225137
YWORDS
URCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 4015)
MORIN, G.B.
TITLE Human telomerase catalytic subunit variants
JOURNAL Patent: US 6337200-A 1 08-JAN-2002;
AUTHORS Location/Qualifiers
SOURCE 1. .4015
/organism="unknown"
SE COUNT 663 a 1363 c 1275 g 714 t
IGIN

```

## Ignment Scores:

```

ed. No.: 4,54e-174 Length: 4015
ore: 3505.00 Matches: 719
cent Similarity: 72.94% Conservative: 122
st Local Similarity: 62.36% Mismatches: 260
ery Match: 59.40% Indels: 52
: 6 Gaps: 13

```

-09-042-460-2 (1-1122) x AR182221 (1-4015)

```

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGGTGCTGCTGCTGCGAGCCACTACCGCGAG 115
QY 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
Db 116 GTGCTGCGCTGGCCACGTTCTGCTGGCGCTGGCGCCCGAGGCTGGCGCTGGTGCGAG 175
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
Db 176 CCGCGGGACCGCGCGCTTTCCGCGCGCTGCTGCGCCAGTGCCTGGTGTGCTGCCCTGG 235
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
Db 236 GAGCAGCGCGCGCGCGCGCTCTCTTCGCCAGGTGTCTCTGCTGAGAGGAGCTG 295
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
Db 296 GTGGCGCGAGTGTCTGAGAGCTGTGCGAGCGCGCGCGAGAACCTGTCTGGCTTCGGC 355
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
Db 356 TTCGGCTGTCTGAGCGCGCGCGCGCGCGCTTACCCAGGCTTACCCAGCGGTGGC 415
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaThrMetLeuLeu 140
Db 416 AGCTACCTGCTCCACCGACCGCTGCGGAGCGCGCGCGCTGGGGCTGCTG 475
QY 141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
Db 476 CTGCGCGCGTGGCGAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
Db 536 CTGTGTGCTCCCGCTGCGCTACAGGTGTGGCGCGCGCGCTGTACAGCTCGCGCT 595
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
Db 596 GCCACTCAGGCGCGCGCGCGCGCGCTAGT---GGACCCGGAAGCGGTCTGGGA--- 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
Db 650 -----TGGAACGGCGCTGGNACCATAGCTCAGGAGCGCGG 688
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
Db 689 GTCCCGCTGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
Db 749 CTGCGCTGCGCCAGAGCGCGCGCGCTGGCGCTGCGCTGCGCGCGAGCGCGCGCGT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
Db 809 GGGAGGCGCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTCTGT 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
Db 869 GTGGTGTCACTTGCAGAG---CCCGCGCGAGAGCGCGCGCGCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 914 GGTGGCTCTCTGGCAGCGCGCGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 974 CCCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGTACGCC 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACTTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370

```



1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 3293 CTCAGAGTACGACACCGTGTACCTACGTGCGACCTCTCTGGGGTCACTCAGGACAGCC 3352  
 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
 3353 CAGACGACACTGAGTCGGAAGCTCCCGGGGACGACGCTGCTGCCCTGGAGCGCGACGC 3412  
 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 3413 AACCCGGCACTGCCCTCAGACTTCAGACCACTCTCTGGAC 3451

SULF 11  
 224455  
 CUS  
 FINITION Sequence 1 from patent US 6440735.  
 CESSION AR224455  
 RSION AR224455.1 GI:23333293  
 WORDS  
 URCE Unknown.  
 ORGANISM Unknown.  
 PERENCE Unclassified.  
 AUTHORS Gaeta, P.C.A.  
 TITLE Dendritic cell vaccine containing telomerase reverse transcriptase for the treatment of cancer  
 JOURNAL Patent: US 6440735-A 1 27-AUG-2002;  
 ATURES Location/Qualifiers  
 1. .4015  
 /organism="unknown"  
 SE COUNT 663 a 1363 c 1275 g 714 t  
 IGIN

ignment Scores:  
 ed. No.: 4,54e-174 Length: 4015  
 ore: 3505.00 Matches: 719  
 cent Similarity: 72.94% Conservative: 122  
 st Local Similarity: 62.36% Mismatches: 260  
 ery Match: 59.40% Indels: 52  
 : 6 Gaps: 13

-09-042-460-2 (1-1122) x AR224455 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCCGGCGCTCCCGCTCCCGACCGCTGCGCTCCCTGCTGCGGACGACCTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
 116 GTGCTCGCGCTGGCCACGTTCTGCGGCGCTGCGGCGCTGCGGCGCTGCGGCTGCTGCG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGCGGGGACCGCGCGCTTTCGCGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTGCGGCG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GAGCGACGCGCGCGCGCGCGCTTCTCCCGCGCTGCTGCGGCGCTGCTGCGGCGCTGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGGCGCGAGTGTGTCAGAGGTGTGTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGCTGGACGGGCG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGCTACCTGCCCCAACACGCTGACCCACGACGCTGCGGGGAGCGGGGCGCTGCGGCGCTG 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160

Db 476 CTGCGCGCGCTGGCGGACGACGCTGCTGCTTCCACCTGCGACGCTGCGCGCTTGTGG 535  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 536 CTGCTGGCTCCAGCTGCGCTTACAGGTGTGCGGCGCGCTGTACAGCTCGCGCT 595  
 QY 181 ThrTrpAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 596 GCCACTAGCGCGCGCGCGCGCGCTAGT---GGACCCGGAAGCGCTCTGGGA--- 649  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
 Db 650 -----TGCGAACGGCGCTGGAACCATAGCTCAGGAGGCGCGG 698  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
 Db 689 GTCCCGCTGGGCTTGCAGCGCGCGGTGCGAGGCGCGGCGCGCTGCGAGCCGAACT 748  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro--- 259  
 Db 749 CTGCGCTTGCACAGAGCGCGCGCTGCGCTGCGCTGCGCTGCGCGCGCGCGCGCGCT 808  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 809 GGGCAGGGGTCTCTGGCGCGCGCGCGCGCGCTGCGACCGAGTGCCTGCTTCTGT 868  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerLys 292  
 Db 869 GTGGTGTCTCCTGCG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 914 GGTGGCTCTCTGGCAGCG 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 974 CCCCCTCCACATCG 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 1034 GAGACCAAGCACTTCT 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 Db 1091 CTACTCAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 Db 1151 CTGGGTCTCCAGCG 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1211 TACTGGCAATGCG 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1271 GGGGTGCTCTCAAGACGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1321  
 QY 431 Leu-----GTACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 432  
 Db 1322 GCGGTGTCTGTGCGCGGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1382 ACAGACCG 1441  
 QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1442 GCGTGTGCG 1501  
 QY 473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492

```

1502 AACGAACGCCGCTTCTCTCAGGAACCAAGAAAGTTTCATCTCCCTCGGGAGCATGCCAAG 1561
493 LeuSerLeuGlnLeuMetTrpTyrMetLysValGluAspCysHisTrpLeuArgSer 512
1562 CTCTCGCTGAGGAGTGAGTGGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGCAGG 1621
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleuAla 532
1622 AGCCAGGGGTGGCTGTGTCTCGGCGCGCAGAGCACCGTCTCGGTGAGAGATCTGGCC 1681
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
1682 AAGTTCCTGCACTGGCTGATGAGTGTGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
1742 GTACGAGAGACCAAGTTCACAAAGAACAGCGCTCTTTCTACCGAAGAGTGTCTGGAGC 1801
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1802 AAGTTCGAAGACATTGGAATCAGACAGCACTTGAAGAGGCTGACGTGCGGAGCTGTGC 1861
593 GlnGluValAlaArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612
1862 GAACGAGAGTTCAGGACGATCGGAAGACGAGCGCGCCCTGTCACGCTCCAGACTCCGC 1921
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1922 TTCTATCCCAAGCTGACGGGCTGCGGGCTGCGGGCATTTGTGACATGACATGCTGCTGG 1981
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1982 AGAAGCTTCGCGAGAGAAAGAGCGGACGCGCTCTCACTCGAGGCTGAAGCACTGTTC 2041
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
2042 AGCTGTCTCAACTACGACGCGGCGCGCGCGCGCTCTCTCGGGCGCTCTGTGTGGCG 2101
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2102 CTGGACCATATCCACAGGCTTGGCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2162 CCGCTGAGCTGACTTGTGTAAGTGGATGTGACGGGCGGTACGACACCATCCCGCAG 2221
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGAACACGTACTGCGTG 2278
733 ArgGlnTyrAlaValAlaArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2279 CGTCGGTATGCGGTGTCAGAAAGCGCGCCATCGGCGACGTCCGCAAGCGCTTCAAGAGC 2338
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2339 CAGTCTCTACTTTCAGACCTTCAGCGGTATGCGGACGTTCGTGCTGCTGCTGCTGCTG 2398
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnGlnSerIleSerMetAsn 792
2399 GAGACC-----AGCCCGCTGAGGATGCGGTGCTCATCGACGACAGCTCTCTCCGTGA 2452
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
2453 GAGCCAGCAGTGGCTCTTCGACGTCTTCCTACGCTTCTTCGTCGACCCAGCCCGCTGCG 2512
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2513 ATCAGGGGCAAGTCTACGTCCAGTGCAGGGATCCGCGAGGCTCCATCTCTCCACG 2572
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2573 CTGCTCTGACCGCTGTCTACGCGGACATGAGAAAGCTGTTTTCGCGGATTCGCGG 2632
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872
2633 GACGGGCTGCTCTCGCTTGGTGATTTCTTGTGTGACACCTCACCTCACCCAC 2692
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2693 GCGAATACTTCTCTCAGACCCCTGGTCGAGGTGTCTCTGAGTATGGCTGGTGGTGAAC 2752
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2753 TTGCGAAGACAGTGTGTAAGTCTCCCTGTAGAACAGAGCCCTGGGTGGCAGCGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATGCGCGGCCACAGGCTATTCCCTGGTGGCGCTGTCTGTCTGTGTATACCGGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGGAGGTGCAGAGGCACTACTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCAC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972
2933 TTCAACCGCGGCTTCAAGCTGGGAGAACATGGCTGCGAAACTCTTTGGGGTCTTCGGG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGCGCTGTTCTGGATTGAGGTGAACAGCTCCAGCGGTGTGCACC 3052
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTCAAGATCTCTCTGTCGAGCGTACAGGTTTACGCACTGTGTGCTGCACTC 3112
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCAATTTCACAGAGTTTGGAGAACCCACATTTTCTGCGCGTCATCTCTGACACG 3172
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCTCTCTCTGTCTACTCTCATCTGAAAGCCAGAACCGAGGATGTGCTGGGGGCCAAG 3232
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
3233 GCGCGCGCGCGCTCTGCTCCGAGCGCGTGCAGTGGCTGCGCCACCAAGCAATTCCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAGCTGACTCGACACCGTGTACCTACGTGCCACTCTCTGGGTCACTCAGGACAGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
3353 CAGACGACGTGAGTCGGAAGCTCCCGGGAGCGAGCTGCTGCTGCTGCTGCTGCTGCTG 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCGGCACTCCCTCAGACTTCAAGACCATCTCTGGAC 3451

RESULT 12
AR226390
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 4015)
AUTHORS
Cech,T.R., Lingner,J., Nakamura,T., Chapman,K.B., Morin,G.B.,
Harley,C.B. and Andrews,W.H.
TITLE
Antisense compositions for detecting and inhibiting telomerase
reverse transcriptase
JOURNAL
Patent: US 6444650-A 1 03-SEP-2002;

```

```

ATUES      Location/Qualifiers
Source      1..4015
            /organism="unknown"
3E COUNT    663 a 1363 c 1275 g 714 t
IGIN
ignment Scores:
ad. No.:    4,54e-174      Length:      4015
pre:        3505.00        Matches:    719
cent Similarity: 72.94%    Conservative: 122
at Local Similarity: 62.36% Mismatches: 260
ary Match:   59.40%      Indels:     52
:            6            Gaps:      13

-09-042-460-2 (1-1122) x AR226390 (1-4015)
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGCGCGCGGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115
21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGCTGCCCTGCCACAGTCTGTGCGCGCCCTGGGCGCCCAAGGCTGCGCGCTGTGTCAG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
176 CGCGGGAGCCCGCGGCTTCCGCGCGCTGTGTGCGCCAGTGCCTGGTGTGCTGCCCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGlyLeu 80
236 GACGACGCGCGCCCGCGCGCCCTTCTTCCGCCAGGTGCTCTGCCGTGAAGGAGCTG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGGCCGAGTCTCAGAGGCTGTGCGAGCGCGCGCGAGACGTCGTGGCTTCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGCTGGACGGGGCGCGCGGGGGCGCGCCCGAGGGCTTTCACCCAGCGTGGC 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
416 AGCTACCTGCCCAACACGGTGNACCGACGACATGCGGGGAGCGGGGCTGGGGCTGTG 475
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGCGCGGCTGGCGGACGACGCTGTGCTTCACTGTGCGCAGCTGCGCGCTCTTTGTG 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGGTGGCTCCAGCTGGCTTACAGGTGTGCGGGCGCGCGCTGTATCCAGCTCGCGCT 595
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCCACTCAGGCGCGCGCGCGCACACGCTAGT---GGACCCCGAAGCGCTCTGGA--- 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
650 -----TGCAACGGGCTGGAAACCATACGTCGAGGAGCGCGG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
689 GTCCCTCTGGGCTTCGACCGCGGCTGGAGAGGCGCGGGGCGAGTCGACCGCGAGT 748
241 ValProSerAlaLysIleAlaArgCysTyrProValProArgValGluGlyPro--- 259
749 CTGCGCTTCCCAAGAGGCCCGCGCGCTGCGCGCTGAGCGGAGCGGAGCGCGCTT 808
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
809 GGGCAGGGGTCTGGGCGCCACCGGGGAGAGCGGTGGACCGAGTGACCGTGTCTGT 868
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292

```

1922 TTCATCCCAAGCCTGACGGGCTGCGCCGATTTGTGAACATGGAATACGTGCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLeuLeuThrLeuPhe 652  
1982 AGACGTTCCCGCAGAGAAAGAGGCGCCAGCGTCTCACTCGAGGGTGAAGGCACATGTTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCAACTACGAGCGCGCGCGCCCGCCCTCTCTGGCGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCAGAGCCTGGCGCCTCTGTGCTGTGCTGTGCGGGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGACTTTGTCAGGTGGATGTGACGGCGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGCTCAGCGAGTCAATCCCGACATCAAAA-----CCCCAGACATGACTGCGTG 2278  
733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGGTATGCCGTGCTCAGAAGCGCGCCATGGGCACGTCCGCAAGGCGCTTCAAGAGC 2338  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACCTTGACAGACTCCAGCGGTACATGCCACAGTTCTGTGGTCACTCGAC 2398  
773 AspSerAspAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGTAGGAGTCCGTCGTCATCGAGCAGAGTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCCAGAGTGGCTCTTCACGCTCTTCATCGCTTCAATGTCACACCGCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTCTCAGGACCTGTCGAGGTGTCCTGAGTATGGTGGTGGTGAAC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGCGCTGTCTACGCGACATGAGAAACAAGCTGTGTCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGTGCTCTGCGTTGGTGGATGATTTCTTGTGGTGAACCTTCACTCACCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTGTCGAGGTGTCCTGAGTATGGTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTCCGGAAGACAGTGTGTAATCTTCCCTGTAGAGACGAGGCGCTGGGGTGGACGGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrPheGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACGCGCTATTTCCCTGCTGGTGGCGCTGCTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGAGAGCGCACTACTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCAC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGAGGAACTGCGTGGCAAACTCTTTGGGGTCTTGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTCAACGCGCTTTCTGGATTTGAGGTGAACAGAGCTCCAGACCGGTGTGCACC 3052

QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGTCGAGCGGTACAGTTTTCAGCATGTGTGTGTCGAGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTTATCAGCAGTTTGGAGAACCACCATTTTCTCGCGCTCATCTCTGACAGC 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCCTCTGCTACTCATCTCTGAAAGCAAGACGAGGGATGTGCTGGGGGCGCAAG 3232  
QY 1053 GlySer-----PheProGluAlaAlaHisIlePheLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCTCTGCTCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGCTGACTCGACACCGGTGTACCTAGTCCACTCTCTGGGGTCACTCAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGCGAGCTGAGTGGAAAGCTCCCGGGGACGAGCTGACTCCCTGGAGCGCGAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGCGACTGCGCTCAGACTTCAAGACCATCTCGAC 3451  
RESULT 13  
AR243328 4015 bp mRNA linear PAT 20-DEC-2002  
LOCUS Sequence 1 from patent US 6475789.  
DEFINITION AR243328  
ACCESSION AR243328  
VERSION AR243328.1 GI:27290539  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 4015)  
AUTHORS Cech,T.R., Lingner,J., Nakamura,T., Chapman,K.B., Morin,G.B.,  
Harley,C.B. and Andrews,W.H.  
TITLE Human telomerase catalytic subunit: diagnostic and therapeutic  
methods  
JOURNAL US 6475789-A 1 05-NOV-2002;  
FEATURES Location/Qualifiers  
source 1..4015  
/organism="unknown"  
BASE COUNT 663 a 1363 c 1275 g 714 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,54e-174 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13  
US-09-042-460-2 (1-1122) x AR243328 (1-4015)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGGCAGCCACTACCGGAG 115  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGCGCACCTGTGTGCGCGCTGCGGGGCCCGAGGCTGCGGCTGGTGAG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGGACCCCGCGGCTTTCCGCGCGCTGGTGCCAGTGGTGGTGGTGGTGGTGG 235

1271	Db	GGGGTGTCCTCTCAAGAGCACATGCCCGCTCGAGCTGGC-----GTCAACCCAGCA	1321
431	Qy	Leu-----Asn 432	
1322	Db	GCGGTGTCTGTGCCGGAGAACCCAGGCGCTGTGTGGCGGCCCCGAGGAGGAGAC	1381
433	Qy	ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr	452
1382	Db	ACAGACCCCGCTCCGCTGGTGCAGCTGCTCCGCGAGCAACAGCAGCCCTGCGAGGTGAC	1441
453	Qy	GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis	472
1442	Db	GGCTTCGTGGCGGCGCTGCTCGCGCGCTGTGTGCCCGCAGGCGCTGTGGGGCTCAGGAC	1501
473	Qy	AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys	492
1502	Db	AACGAACCGCTCTCTCAGGAACACCCAGAGATTATCTCCCTGGGAAGCATGCCAAG	1561
493	Qy	LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer	512
1562	Db	CNCTCGCTGCAGAGCTGACGTGGAGATGAGCGTGGGATCTGCGCTGTGCTGCGCAGG	1621
513	Qy	SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla	532
1622	Db	AGCCACGGGTGGCTGTGTTCGGCGCGAGAGCACCGTCTGCGTGAGAGATCTCTGCC	1681
533	Qy	ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr	552
1682	Db	AAGTTCCTGCATCGCTGATGAGTGTGTAGTCTCGAGCTGCTCAGGTCTTCTTTTAT	1741
553	Qy	IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer	572
1742	Db	GTACGGGAGACCATGTTTCAAGAAGACAGGCTCTTTTCTACCGAAGAGTGTCTGAGC	1801
573	Qy	LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer	592
1802	Db	AAGTTGCAAGCATTTGAAATGCACACAGCACATGGAAGGGTGCAGCTCGCGAGCTGCG	1861
593	Qy	GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg	612
1862	Db	GAGCAGAGGTTCAGGAGCATTCGGAGACCGCGCCCTGCTGAGCTCCAGACTCTCGC	1921
613	Qy	PheIleProLysProAsnGlyLeuArgProIleValAlaAsnMetSerTyrSerMetGlyThr	632
1922	Db	TTATCATCCCAAGCTGACGGCTGCGCGCATTTGTGTAAATCATGACTACGTGCTGGAGCC	1981
633	Qy	ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe	652
1982	Db	AGACGTTCCGACAGAAAAGAGCGCGCGTCACTCACCCTCGAGGGTGAAGCGACTGTTC	2041
653	Qy	SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly	672
2042	Db	AGCGTGTCAACTACGAGCGGCGCGCGCCCGCGCTCTCTGGGCGCTCTGTCTGGCG	2101
673	Qy	MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln	692
2102	Db	CTGACGATATTCACAGGGCTCGCGCATCTTCTGTCGTGCTGCGGGCCCGACGACCCG	2161
693	Qy	ThrProArgMetTyrPheValLysAlaAspValThrGlyValaTyrAspAlaIleProGln	712
2162	Db	CGCCTGTAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTGACACACCATCCCCAG	2221
713	Qy	GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle	732
2222	Db	GACAGGCTCAGCGAGGTTCATCGCAGCATCAACAA---CCCCAGAAACACGTACTGGGTG	2278
733	Qy	ArgGlnTyrAlaValValArgArgAspSerGlnGlnValHisLysSerPheArgArg	752
2279	Db	CGTCGGTATCGGTTCAGAGAGCGGCCCATGTGGCACGTCCGCAAGGCCTTCAGAGC	2338
753	Qy	GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln	772





550 -----TCCGAACGGGCTCGAACCATAGCGTCAGGAGCGCCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCTCGGCGCTGCAGCCCGGCTGCAGAGCGCGGGCGAGTCCAGCGCAAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluPro--- 259  
749 CTGCCGTTCGCCAAGAGCCAGCGCTGCGGCTGCGCCCTCAGCCGAGCGAGCGCCGTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGGCTCTGGGCGCCACCGCGGCGAGCGCGTGGACCGAGTACCGTGTTCGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTACCTCCAG-----CCGCGCGAAGAACCCACCTCTTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTTGGCTCTCTGCAGCGGCCACTCCACCCATCGTGGCGCGCCAGCACCGCGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCATCGCGGCCACACACTCCCTCGGCACACGCCCTTGTCCCGGTGTACGCC 1033  
331 GlyThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTTACTCTCTCAGGGAC---AAGGACAGCTGGGGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCGAGGCCAGCCCTGACTGGCGCTCGGAGGCTCGTGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTCAGGCGCTCGAGCGCAGGAGACTCCCGCAGGCTTCCCGCGCTGCCCGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTGTCTTCGGAGCTGCTTGGGAACACACAGCGAGTCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGCGTGTCTCTCAAGACGCACTCCCGCTGCGAGCTGGG-----GTCAACCCCA 1321  
431 Leu-----Asn 432  
1322 GCCGCTCTGTCCCGGAGAGCCCGAGGCTCTGTGGCGGCCCGGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCTGTGCAGCTGCTCCCGCAGCACAGCAGCCCTTGGCAGGTGAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGTGGCGGCTGCTTCCCGCGCGCTGGTGGCCCGCAGGCTCTGGGGCTCAGGCGC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492  
1502 AACGAACCGCGCTCTCTCAGGAACCAAGAGTTTCATCTCCCTCGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTCAGGAGTGCAGTGAAGATGACGCTGGCGGACTTGGCTGTGGCTGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGGGTGGCTGTTCGGCGCGCAGAGACCGCTCTCGGTGAGAGATCCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGTGTAGTGTACGTCTGCTGAGCTGCTGAGGCTGCTGAGTCTTTT 1741

553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTCACGGAGACACACCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGGATTTGGATCAGACAGCACTTGAAGAGGTGACAGTGGGAGCTGTGC 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCGCCCTGTCTGAGCTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTATCCCCCAAGCTGAGCGGCTGCGCGCATTTGAACATGGACTAGCTGCTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGGCGCCGAGGCTCTACCTCGAGGGTGAAGGCACTGTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCACTACAGCGGCGCGCGCCCGGCTCTCTGGCGGCTCTGTCTGCTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGCGCTGGCGCACCTTCGTGCTGTGCGGGCCCGACGCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCCCTGAGCTTACTTTGTCAAGGTGATGTGACGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGCTCAGGAGTTCATCCGACGATCATCAA---CCCCAGAACACACTCTGGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCGCTGTCAGAAAGCGCCCATGGCACGTCGCGAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGluProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTTACCTTACAGACCTCCAGCGGTACATGCGACAGTTCGTGGGTCACTCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGAGTCCGCTCATCGAGCAGACTCCTCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
2453 GAGGCGCAGCTGGCTCTTCGACGCTTCCTACGCTTCATGTGCCACCGCGCGTGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTCAGTCCAGTGGCAGGGATCCGCGAGGCTCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCGAGCTGTGTACGGGACATGAGAGAACAAAGCTGTTTGGGGGATTCGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GAGGGCTGCTCTCTGGTGTGATGATTTCTGTTGTGACACCTCCTCCTCACCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTTGTCCGAGGTGTCTCTGAGTATGCTGCTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912  
2753 TTGCGGAACAGAGTGGTGAACCTTCCCTGTAGAACGAGCGCCCTGGGTGGACCGGCTTT 2812

```

913  TyGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATGCGCGCCACCGCCCTATTCCTCGTGTGGCGCTGCTGCGATACCGCGACC 2872
933  LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGGAGGTGCAGAGCGACTACTCCAGTATGCCCGACCTCCATCAGACCGAGTCTCACC 2932
953  PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2933 TTCAACCGCGCTTCAAGGTGGAGGAACATCGTCGCAAACTCTTTGGGGTCTGGCG 2992
973  LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGCCCTGTTCTGGATTTCAGGTGAACAGCTCCAGACGCTGTCACC 3052
993  AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAGATCTCTCTGCTGCGAGCGGTACAGGTTTCAGCGATGTGTGTCAGCTC 3112
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGGTCTCTCTGACAGC 3172
1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3173 GCCTCCCTCTACTCATCTCTGAAAGCCAGAACGACGAGGATGCTGGGGGCCAAG 3232
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGCCCGCGCGCCCTGCGCTCCGAGCGCGTGCAGTGTGCGCAGCAGCATCTCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAAGCTGACTGCACACCGTGTACCTAGTCCACTCTCGGGTCACTCAGACAGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
3353 CAGACGAGCTGAGTCGAGAGCTCCCGGGACAGACGCTGACTGCCCTGGAGCCGACCC 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCGGCACTGCCCTCAGACTTCAGACCATCTCTGGAC 3451

```

```

RESULT 15
DCUS      AR265996      4015 bp      DNA      linear      PAT 10-APR-2003
SEQUENCE 3 from patent US 6492171.
SEQUENCE  AR265996
SYNOPSIS  AR265996.1  GI:29694842
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 4015)
AUTHORS    Monia,B.P., Gaarde,W.A., Freier,S.M. and Wanciewicz,E.
TITLE       Antisense modulation of TERT expression
JOURNAL     Patent: US 6492171-A 3 10-DEC-2002;
FEATURES    Location/Qualifiers
             1..4015
             /organism="unknown"
BASE COUNT 663 a 1363 c 1275 g 714 t
IGIN
Alignment Scores:
Ref. No.: 4.54e-174      Length: 4015
Score: 3505.00        Matches: 719
Percent Similarity: 72.94%      Conservative: 122
1st Local Similarity: 62.36%      Mismatches: 260
Very Match: 59.40%      Indels: 52
3: 6      Gaps: 13
3-09-042-460-2 (1-1122) x AR265996 (1-4015)

```

```

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
DB 56 ATGCGCGCGGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGACCGACCTACCGCGAG 115
QY 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
DB 116 GTGTGCTGCGCTGCGACAGCTTCTGCGCGCGCTGGGGCCCGACGGCTGGCGGCTGTCAG 175
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
DB 176 CGCGGGGACCCCGCGCGCTTTCGCGCGCTGCTGCGCCACAGTCCCTGCTGCGTCCCTCG 235
QY 61 GlySerProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
DB 236 GAGCAGCGCGCGCGCGCTCCCTCCCTCCCGCAGGTGCTGCTGCTGAAGAGAGCTG 295
QY 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
DB 296 GTGGCCCGAGTCTGTCAGAGGCTGTGCGAGCGCGCGCGAAGACGTCGCTGCGCTCGGC 355
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
DB 356 TTGCGCTCTGTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTG 415
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
DB 416 AGTACTCTCCCAACACAGGTGACCGACGACTGCGGGGCGAGCGGGGCGTGGCGCTGCTG 475
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
DB 476 CTGCGCGCGCTGCGCGAGCAGCTGCTGTTTACCTGCTGGCAGCGCTGCGCGCTCTTTGTG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
DB 536 CTGTGTGCTCCAGCTGCGCTTACAGGTGTGGCGCGCGCGCTGTACCGCTCGCGCT 595
QY 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
DB 596 GCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGluAlaPro 220
DB 650 -----TGCAGCGCGCTGGAACCATAGCTCAGGAGCGCGCGCGCG 688
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
DB 689 GTCCCGCTGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748
QY 241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
DB 749 CTGCGCTTCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
DB 809 GGCGAGGGGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
DB 869 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
DB 914 GGTGCGCTCTTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 974 CCCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyValAspGlyGlnGluArgLeuAsnProSerPhe 350
DB 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGCGAC---AAGGAGAGAGCTGCGCGCGCTCTTC 1090

```



3233 GGCGCCGCCGCCCTCTGCCCTCCGAGGCCGTGTCAGTGGCTGTGCCACCAAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGGTGTACCTACGTGCCACTCCTGGGGTCACTCAGGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACCTGACTGCCCTGGAGCCGCGAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 ACCCGCACTGCCCTCAGACTTCAAGACCAATCCCTGGAC 3451

arch completed: February 1, 2004, 18:05:14  
b time : 10076 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

on: February 1, 2004, 15:05:58 ; Search time 663 Seconds  
(without alignments)  
4568.280 Million cell updates/sec

tle: US-09-042-460-2  
irect score: 5901  
quence: 1 MTRAPCPAVRSLLRSRYRE.....TILKAADPALSTDFQITILD 1122

oring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Xgapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

arched: 2552756 seqs, 1349719017 residues

tal number of hits satisfying chosen parameters: 5105512

nimum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

mmand line parameters:

ODEL=frame+ p2n.model -DEV=xlh  
=/cgn2.1/USPTO.spool/US09042460/runat 01022004.150550.17004/app query.fasta\_1.1287  
B-N Geneseq 19Jun03 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0  
COPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
IST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
QDE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
SER=US09042460 @CGN 1.1.0 @runat 01022004.150550.17004 -NCPU=6 -ICPU=3  
O MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
EV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
GAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

tabase : N Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3513	59.5	4015	20	AAZ08150 Human telomerase r
2	3505	59.4	3396	22	AAH44366 Human telomerase n
3	3505	59.4	3396	22	AAH48235 Heart muscle cell
4	3505	59.4	3396	22	AAH49601 Human coding seque
5	3505	59.4	3399	24	ABZ35720 Human telomerase re
6	3505	59.4	3399	24	ABX09963 Human telomerase r
7	3505	59.4	3399	24	ABV78144 Human telomerase r
8	3505	59.4	3399	24	ABL91685 Human polynucleoti
9	3505	59.4	3798	19	AAV27876 Human telomerase p
10	3505	59.4	3955	19	AAV22379 Human telomerase r
11	3505	59.4	4015	20	AAZ30154 cDNA encoding a hu
12	3505	59.4	4015	20	AAZ20279 Human telomerase r
13	3505	59.4	4015	20	AAZ00724 Human telomerase c
14	3505	59.4	4015	22	AAH45901 Human hTERT gene.
15	3505	59.4	4015	24	AAH46821 Human telomerase r
16	3505	59.4	4015	24	ABA97534 Cancer cell discr
17	3505	59.4	4015	25	ABZ22474 Human telomerase r
18	3505	59.4	4015	25	ABZ18391 Group iii cDNA can
19	3505	59.4	4042	20	AAV72117 Human catalytic te
20	3505	59.4	4070	24	ABL53711 Human telomerase c
21	3498	59.3	3396	20	AAH18266 Telomerase coding
22	3498	59.3	3964	20	AAH18254 Human telomerase c
23	3498	59.3	4023	19	AAV60320 Human telomerase g
24	3498	59.3	4027	20	AAH99424 Human EST2 coding
25	3498	59.3	4027	21	AAA29388 hEST2, a human tel
26	3496	59.2	3453	25	AAV76217 Human TERT coding
27	3496	59.2	4037	19	AAV22428 Human telomerase r
28	3496	59.2	13766	24	AAH46790 pGRN145 plasmid DN
29	3473	58.9	8742	20	AAH18273 PWG5a plasmid DNA
30	3473	58.9	3500	20	AAH18275 Telomerase coding
31	3435	58.2	3918	20	AAH18269 Telomerase coding
32	3435	58.2	3918	20	AAH18278 Altered C-terminus
33	3304.5	56.0	3203	19	AAH18268 Human telomerase r
34	3276.5	55.5	3855	19	AAV22382 Altered C-terminus
35	3275.5	55.5	3323	20	AAH18277 Altered C-terminus
36	3234.5	54.8	3167	20	AAH18271 Altered C-terminus
37	3234.5	54.8	3167	20	AAH18280 Altered C-terminus
38	3141	53.2	7688	20	AAH18351 Telomerase coding
39	3038	51.5	7797	20	AAH18350 Telomerase coding
40	2881.5	48.8	3069	20	AAH18267 Truncated telomera
41	2877.5	48.8	3069	20	AAH18276 Truncated telomera
42	2814.5	47.7	3033	20	AAH18270 Truncated telomera
43	2814.5	47.7	3033	20	AAH18279 Truncated telomera
44	2808.5	47.6	7615	20	AAH18349 Telomerase coding
45	2790.5	47.3	2848	19	AAV27872 Human telomerase p

## ALIGNMENTS

### RESULT 1

AAZ08150  
ID AAZ08150 standard; cDNA; 4015 BP.

XX AC

XX AAZ08150;

XX 17-JAN-2000 (first entry)

XX DE Human telomerase reverse transcriptase cDNA.

XX KW Human telomerase reverse transcriptase; hTERT; telomerase; hEST2;

XX KW catalytic protein component; cell proliferative capacity;

XX KW cell immortality; neoplastic phenotype; diagnostic application;

XX KW prognostic application; telomerase related condition; cancer;

XX KW therapeutic agent; telomerase expression; telomerase activity; ds.

XX OS Homo sapiens.

Key Location/Qualifiers  
 CDS 56..3454  
 /tag= a  
 /product= "Human telomerase reverse transcriptase"  
 /transl\_except= (pos:1877..1879, aa:Gln)

W09950279-A1.

07-OCT-1999.

31-MAR-1999; 99WO-US07160.

31-MAR-1998; 98US-0052919.

(GERO-) GERON CORP.  
 (UYTE-) UNIV TECHNOLOGY CORP.Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB;  
 Andrews WH;

WPI; 1999-610834/52.

P-PSDB; AAY28881.

Antisense polynucleotides for human telomerase reverse transcriptase  
 used for diagnosing or treating cancer -

Claim 1; Fig 1; 31pp; English.

The present sequence encodes for human telomerase reverse transcriptase (hTERT). This is the catalytic protein component of telomerase and is also referred to as hEST2. This correlates with cell proliferative capacity, cell immortality, and the development of a neoplastic phenotype. Human TTT antisense oligonucleotides are useful for diagnostic or prognostic applications to telomerase related conditions, including cancer. They are also useful as therapeutic agents, for inhibition of telomerase expression and activity.

Sequence 4015 BP; 663 A; 1364 C; 1274 G; 714 T; 0 other;

## ignment Scores:

ed. No.:	3,32e-234	Length:	4015
ore:	3513.00	Matches:	720
recent Similarity:	73.03%	Conservative:	122
st Local Similarity:	62.45%	Mismatches:	259
ery Match:	59.53%	Indels:	52
:	20	Gaps:	13

-09-042-460-2 (1-1122) x AAZ08150 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCGGGCGCTCCCGCTCCGAGCGGTGGCTCCCTGCTCGGAGCCACTACCGGAG 115  
 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgValGln 40  
 116 GTGTCGCGCTGCGCAGTTCGCGCGCCCTGCGGCCCTGCGGCTGGCGGTGGTGCG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGCGGGACCCCGCGGCTTCCGCGCGCTGTGTGGCCAGTGCCTGTGTGCGTCCCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GAGCAGCG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGCCCCAGTGTGTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
 101 PheGluLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGTGGACGGGGCCCGCGGGGGCCCGCGGGCGCGCGCGCGCGCGCGCGCG 415

QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 DB 416 AGCTACCTGCCCCAACACGCTGACCGACGACGCGGGGAGCGGGCGGTGGGGCTGCTG 475  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 DB 476 CTGGCGCGCTGGGCGACGACGCTGCTGCTTCACTGCTGGCAGCTGGCGCTCTTTTGG 535  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 DB 536 CTGGTGGCTCCAGCTCGGCTACAGGTGTGGGGCGCGCGCTGTACAGCTCGCGCT 595  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 DB 596 GCCACTCAGGCG 649  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGlnAlaPro 220  
 DB 650 -----TGCMAACGGGCTCGAACCATAGCTCAGGAGGCGCGG 698  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 DB 689 GTCCCCCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro--- 259  
 DB 749 CTGCGTGTGCCAAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 808  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 DB 809 GGGCAGGCGTCTCGGCG 868  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
 DB 869 GTGGTGTACCTGCGAGA-----CCGCGCGAAGACCCACCTTTTGGAG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 914 GGTGGCTCTCTGGCAGCGCGCACTCCCGACCCATCCGTGGGCGCGCGCGCGCGCGCG 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 974 CCCCCATCCATCG 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 1034 GAGACCAAGCACTTCTCTACTCTCTCAGGCGAC---AAGGAGCAGCTGCGCGCTCCTTC 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIleIlePhe 370  
 DB 1091 CTACTCAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 DB 1151 CTGGGTTCCAGCGCGCTGATCCAGGAGCTCCCGCAGGTTCGCCCGCTGCCCGCGCG 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1211 TACTGGCAATGCGCGCGCTGTTCTGAGAGTGTGTGGAAACACCGCGCGAGTGGCCCTAC 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 DB 1271 GGGGTGCTCTCAAGACCACTGCCCGCTGGAGCTGG-----GTACCCCGAGCA 1321  
 QY 431 Leu-----Asn 432  
 DB 1322 GCCGGTGTCTGGCCCGGAGAGCCCGCGCGCTCTGTGGCGCGCGCGCGCGCGCGAGG 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 DB 1382 ACAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441



453 GlyPheLeuArgAlaCysLeuCysLeuValValSerAlaSerLeuThrArgHis 472  
1442 GGCTTCGTGGGGCTGCTGCTGCGCGGTGGTGGCCCGAGGCTCTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLeuLeuValValSerLeuGlyLysThrGlyLys 492  
1502 AACGAACCGCGCTTCCTCAGGACACCAAGAAAGTTTCATCTCCCTCGGGAAGCATCCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGCGAGCTGAGCTGGAAGATGAGCTGCGGATCTGCGTGGCTGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgLysLeuAla 532  
1622 AGCCACGGGTGGCTGCTGCTGCGCGCGAGACACCGCTGCTGCTGAGAGATCTCGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspTrpValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AGTTCTCTGACCTGGCTGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGGAGACACCGTTTCAAGAACACGCTCTTTTCTACCGGAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAGAGCATTTGGAATTCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGCG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GNACAGAGTTCAGGCACCATCGGAGACGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCTGACGGCTGCGCGCTGCGCGATTTGTAACATGACTACGTCTGCGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGTTCGCGAGAGAAAGAGCGCGGAGCGCTCCTCCTCAGGAGTGAAGCATCTGTTC 2041  
653 SerMetLeuAsnTrpGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCAACTACGAGCGGCGCGCGCGCTCTCTGCGCGCTCTGCTGCTGCGG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACCATATCCACAGGGCTCGGCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2221  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCAGCATCATCAAA---CCCCAGAACACGTPACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGCTCTACCTTTCAGACCTCCAGCGGTATCATCGCAGTTCGTGGCTCATCTGCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGGATGCGCTGCTCATCGACAGAGTCTCCCTCGAAT 2452  
793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCCAGAGTGGCTTCTTCGAGCTTCTTCCTACGCTTCTGCTGCTGCTGCTGCTGCTGCT 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832

Db 2513 ATCAGGGGCAAGTCTACGTCAGTGCAGGGATCCGCGAGGCTCCATCTCTCCACG 2572  
Qy 833 LeuLeuCysSerLeuCyPheGlyAspMetGluLeuLysLeuPheAlaGluValGlnArg 852  
Db 2573 CTGCTCTGAGGCTGCTGCTACGCGCAGCATGAGAAACAAGCTGTTTGGGGGATTCGGCGG 2632  
Qy 853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
Db 2633 GACGGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2692  
Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
Db 2693 GCGAAACCTTCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2752  
Qy 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyValAlaPro 912  
Db 2753 TTGCGAAGACAGTGGTGAATTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2812  
Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2813 GTTCAGATCGCGGCCACCGGCTATTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2872  
Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2873 CTGAGGTGCAGAGCGACTTCTCCAGCTATGCCCGACCTCCATCAGAGCCAGTCTCACC 2932  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2933 TTCAACCGCGGCTTCAAGGCTGGAGAGACATGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCT 2992  
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2993 CTGAGTGTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3052  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTCAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3112  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTGCGGCTCATCTCTGACACG 3172  
Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3232  
Qy 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3292  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGCTGACTCGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3352  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACCGCTGAGTGCAGAGCTCCCGGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3412  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGGCACTCCCTCAGACTTCAAGACCATCTCTGGAC 3451

## RESULT 2

AAH44366  
ID AAH44366 standard; cDNA; 3396 BP.  
XX

AC AAH44366;

DT 26-SEP-2001 (first entry)

XX Human telomerase nucleotide sequence SEQ ID NO:32.

DE

XX

1 Differentiation; heart muscle cell; cytokine; transcription factor;  
 2 proliferation; surface antigen; heart disease; cardiomyocyte;  
 3 bone marrow; umbilical blood cell; heart muscle degeneration;  
 4 myocardial infarction; ss.

5 Homo sapiens.

6 WO200148150-A1.

7 05-JUL-2001.

8 02-NOV-2000; 2000WO-JP07741.

9 28-DEC-1999; 93JP-0372826.

10 28-FEB-2000; 2000WO-JP01148.

11 (KYOW) KYOWA HAKKO KOGYO KK.

12 Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

13 Yamada Y;

14 WPI; 2001-425655/45.

15 P-PSDB; AAB99930.

16 Cells capable of differentiating into cardiomyocytes and originating in  
 17 bone marrow or umbilical blood cells for study of cardiomyocyte  
 18 differentiation and treatment of heart disease -

19 Disclosure; Page 141-147; 187pp; Japanese.

20 The present invention describes cells originating in bone marrow or  
 21 umbilical blood cells which are capable of differentiating into  
 22 cardiomyocytes. Also described are: (1) cardiomyocytes produced by the  
 23 differentiation of the cells; (2) a method for carrying out the  
 24 differentiation into cardiomyocytes, regulated by a promotional and/or  
 25 inhibitory factor; (3) a method for the differentiation of the cells  
 26 into cell types other than cardiomyocytes; (4) drug compositions  
 27 promoting the formation of heart muscle and regeneration of heart tissue  
 28 which contain the cells; (5) a method for the production of antibodies  
 29 which recognise the cells, especially antibodies which recognise a  
 30 surface antigen on the cells; (6) a method for screening factors which  
 31 promote the proliferation of the cells; (7) a method for immortalising  
 32 the cells by expressing telomerase in them; (8) drug compositions for  
 33 the treatment of heart disease which contain the immortalised cells; and  
 34 (9) cell-free supernatant from the culture of the cells and its use in  
 35 promoting their differentiation into cardiomyocytes. The cells are used  
 36 in the treatment of diseases involving heart muscle degeneration, such  
 37 as myocardial infarction and in the study of cardiomyocyte  
 38 differentiation. AAH4351 to AAH4409 and AAB99915 to AAB99935 represent  
 39 sequences used in the exemplification of the present invention.

40 Q Sequence 3396 BP; 549 A; 1157 C; 1088 G; 602 T; 0 other;

41 Alignment Scores:

red. No.:	9,49e-234	Length:	3396
core:	3505.00	Matches:	719
Percent Similarity:	72.94%	Conservative:	122
Best Local Similarity:	62.36%	Mismatches:	260
Query Match:	59.40%	Indels:	52
B:	22	Gaps:	13

42 S-09-042-460-2 (1-1122) x AAH44366 (1-3396)

43 Y	1	MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu	20
44 b	1	ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG	60
45 Y	21	ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln	40
46 b	61	GTGTCCTCCCTGCGCAGCTTGTGCGCGCTTGGGCGCCCGAGGCTGGCGCTGTGTGAG	120
47 Y	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr	60

121	CGCGGGGACCCGCGGCTTCCGCGGCTGTGTGGCCAGTGCCTGCTGTGTGCTGCTGCTG	180
61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
181	GAGCAGCGGCGCGCGCGCGCTCTTCCGCCAGGTGCTGCTGCTGCTGCTGCTGCTG	240
81	ValAlaArgValValGlnArgCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
241	GTGCGCGAGTGTGTGAGAGGCTGTGCGAGCGCGCGCGGAGAACGTGTGCTGCTGCGC	300
101	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120
301	TTGCGCGTGTGCGAGC	360
121	SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValAlaTyrMetLeu	140
361	AGTACTCTGCCCAACACGCTGACCGACGACCTGCGCGCGCGCGCGCGCGCGCGCGC	420
141	LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu	160
421	CTGCGCGCTGTGCGCGACGACGCTGTGTGTACCTGCTGGCAGCGCTGCTGCTGCTG	480
161	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla	180
481	CTGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGC	540
181	ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg	200
541	GCACCTCAGC	594
201	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro	220
595	-----TGCAGAACCGCGCTGGAACCATAGCGCTCAGGAGCGCGCGG	633
221	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer	240
634	GTCCCCCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGC	693
241	ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro---	259
694	CTGCGCTGCT	753
260	-----HisArgGlnValLeuProThrProSerGlyLysSerTrp---	272
754	GGCGAGCGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGT	813
273	ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys	292
814	GTGCTGTCACTTCCCGA-----CCCGCGGAGAGCGCGCGCGCGCGCGCGCGCGC	858
293	GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer	311
859	GGTGGCTCTCTGCGCAGC	918
312	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle	330
919	CCCCATCCACATCGC	978
331	GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe	350
979	GAGACCAAGCACCTTCTCTACTCTCTCAGCGCGA---AAGGAGCAGCTGCGCGCGCTTC	1035
351	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeuPhe	370
1036	CTACTAGCTCTCTGAGGCCACCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTTC	1095
371	LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg	390
1096	CTGGTTCAGGCCCTGAGTCCAGGAGCTCCCGCGAGGTGCGCGCGCTGCGCGCGCGC	1155
391	TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr	410
1156	TACTGGCAAAATGCGCGCGCTGTTTCTGGAGCTGCTTGGGAGACCGCGCGCTGCTGCT	1215

```
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1216 GGGGTGCTCTCAAGACGACTGCCCGCTGCAGGTGGG-----GTCACCCCAGCA 1266
431 Leu-----Asn 432
1267 GCCGGTGTCTGTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrrpGlnValTyr 452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1327 ACAGACCCCGCTGCTGTGAGCTGTCTCCGACAGCACAGCAGCCCTTGGCAGGTGAC 1386
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrrpGlyThrArgHis 472
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1387 GGTCTGTGGGGGCTGCTGTGGCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1446
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1447 AACGAAGCGGCTTCTCAGGAACACCAAGAGTTCACTCTCCTGGGGAAGCATGCCAAG 1506
493 LeuSerLeuGlnLeuMetTrrpLysMetLysValGluAspCysHisTrrpLeuArgSer 512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1507 CTCTCGCTGCAGAGCTGACGTGGAAGATGAGCGTGCAGGACTGCGCTTGGCTGCGCAGG 1566
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1567 AGCCAGGGGTGGCTGTCTCCGCGCGCAGACACCGTCTGCGTGAAGAGATCTTGCC 1626
533 ThrPheLeuPheTrrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1627 AAGTTCTCTGCACTGGTGTAGTGTGTACGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1686
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrrpSer 572
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1687 GTCCAGAGACCAAGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC 1746
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1747 AAGTTGCAAGACATTGGATCATCAGACAGCACTTGAAGAGGTGCAGCTCGCGGAGCTGTC 1806
593 GlnGluGluValArgHisGlnAspThrTrrpLeuAlaMetProIleCysArgLeuArg 612
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1807 GAAGCAGAGAGTCAAGCAGCATCGGAGAGCAGCGCCCGCTCTGACGTCCAGACATCCGC 1866
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1867 TTCAATCCCAAGCTCAGCGGCTGCGGCGGATGTGAACATGGACTAGCTCGTGGAGACC 1926
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1927 AGAACGTTCCGCAGAGAAAAGAGGCGGAGCGCTCTCAGCTCGAGGGTGAAGGCATGTTTC 1986
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1987 AGCGTGTCTCACTACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2046
673 MetAsnAspIleTyrArgThrTrrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2047 CTGGACGATATCCACAGGCGCTGGGCGACCTTCGTGTGCTGTGCTGTGCTGTGCTGTGCT 2106
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2107 CCGCCTGAGCTGTACTTGTCAAGTGGATGTGAGGCGCGGTACGACACCATCCCCCAG 2166
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2167 GACAGGCTCAGCGAGTCTATCCGACGATCATCAAA-----CCCCAGAACATCTGGGTG 2223
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2224 CGTGGTATGCGGTGTCTCCAGAAAGCCGCCATGGGCGACGCTCCGCAAGGCGCTTCAAGAGC 2283
```

```
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2284 CACGCTCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTTCTGGTGTCACTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2344 GAGACC-----AGCCCGCTGAGGAGTCCGCTCTCATCGACGAGCTCTCTCCCTGAAT 2397
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2398 GAGGCGCAGCAGTGGGCTCTTCGACGTCTTCCTACGCTTCATGTGCCACACCGCGTGCAG 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyLeuProGlnGlySerSerLeuSerThr 832
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2458 ATCAGGGCGAAGTCTCTGCTCAGTTCAGGGGATCCCGAGGGCTCCATCTCTCCACG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2518 CTGCTCTGCAGCCTGTGCTACGGCGCATGAGAGAAAGCTGTTTGGGGGATTCGGCGG 2577
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2578 GACGGGCTGCTCTGCGTTCGTTGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2638 GCGAAAACCTTCTCAGGACCTGCTCGAGGTGCTCCTGAGTATGCTGGTGGTGAAC 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2698 TTTCGGAAGACAGTGTGAACCTCCCTGTAGAACGAGGCCCTCGGTGGTGGCAGCGCTTT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTrrpCysGlyLeuLeuLeuAspThrGlnThr 932
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2758 GTTCAGATGCCGCGCCACCGGCTATTCCCTGTGTGGGCTGTGTGTGTGTGTGTGTGT 2817
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2818 CTGAGGTGTGCAGAGCGACTACTCCAGTATGCCCGACCTCCATCAGAGCGAGTCTCACC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2878 TTCAACCGCGCTCAAGGCTGGAGAACATCGCTCGCAAACTTTTGGGGTCTTGGCG 2937
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2938 CTGAGTGTACAGCTGCTTCTCGATTTCAGGTGACAGCTCCAGCGGTGTGCACC 2997
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2998 AACATCTCAAGATCCTCTGCTGAGCGGTACAGGTTTCACGCATGTGTGTGTGTGTGT 3057
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3058 CCATTTTCATCAGCAAGTTTGAAGAACCCACATTTTCTGGCGGTCTCTGTACAGC 3117
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3118 GCTCTCTCTCTACTCATCTCTGAGCGGTACAGGTTTCACGCATGTGTGTGTGTGTGTGT 3177
1053 GlySer-----PheProGluAlaAlaHisTrrpLeuCysTyrGlnAlaPheLeu 1069
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3178 GCGCGCGCGCGGCTCTGCGCTCCGAGCGCGTGCAGTGGTGTGTGCCAACCAAGCATTCCTG 3237
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3238 CTCAGCTGACTGCAGACCGGTGTCACTACGTGCCATCTCTGGGGTCACTCAGGACAGCC 3297
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3298 CAGACGCACTGAGTGGAAAGCTCCCGGGACGACGCTGACTCCTCCCTGGAGCGCGAGCC 3357
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
```

3358 AACCCGGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3396

STH.T 3

36645  
5 1708

H48235  
AAH48235 standard; DNA: 3396 BP.

AAH48235:

21-SEP-2001 (first entry)

Heart muscle cell differentiation related DNA SEQ ID NO: 32.

Heart muscle cell; human; cell differentiation; heart disease; ds.

*Homo sapiens.*

WO200148151-A1.

05-JUL-2001.

27-DEC-2000: 2000WO-JP09323.

28-DEC-1999: 99JP-0372826.

28-DEC-1999; 2000F-03/2020;  
28-FEB-2000; 2000WO-JP01148;

02-NOV-2000; 2000WO-JP07741.

(KYOW ) KYOWA HAKKO KOGYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

**Yamada Y;**

WPI; 2001-425656/45.

P-PSDB; AAG64859.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease -

Disclosure; Page 147-153; 183pp; Japanese.

The present invention provides cells originating in the human bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. These cells are useful in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction, and the study of cardiomyocyte differentiation. The present sequence is an oligonucleotide described in the exemplification of the invention.

Sequence 3396 BP; 549 A; 1157 C; 1088 G; 602 T; 0 other;

Assignment Scores:

ed. No.:

Score: 3505.00

Percent Sim

### Best Local

## Query Match

33

○ ○ ○ ○ ○

1-09-042-460-2 (1-1122) x AAH48235 (1-3396)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20

1 ATGCCGGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGTGCGAGCCACTACCGCGAG 60

21 ValTrpProLeuAlaThrPheValArqArqLeuGlyProGluValArqArqLeuValGln 40

61 GTGCTGCCGTGGCCACGTTCTGTCGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCGAG 120

41 proGlyAspProLysIleTyrArgqThrLeuValAlaGlnCysLeuValCysMethHisTrp 60

121 CGCGGGACCGCGGCTTTCCGCGCGCTGGTGCCAGTCCTGTGTGCGTGCCCTGG 180

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

1216 GGGGTGCTCTCTCAAGACGCACTGCCCGCTGCGAGCTGCG-----GTCAACCCAGCA 1266  
431 Leu-----Asn 432  
1267 GCGGTGCTGTGTCGCCGGAGAGAGCCAGGCTCTGTGGCGCCCGAGGAGGAGC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrTrpGlnValTyr 452  
1327 ACAGACCCCGCTGCTGTGTGAGCTGTCTCCGCGACAGCAGCGCCCTGCGAGGTATC 1386  
453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTyrGlyThrArgHis 472  
1387 GGTTCGTGCGGCGCTGCTGCGCGCTGTGTGCGCCCGAGGCTCTGGGGCTCCAGGCAC 1446  
473 AsnGluArgPhePheLeuLeuLeuValPheLeuLeuGlyValTyrGlyValTyr 492  
1447 AACCAACGCGCTTCTCAGGAACACCAAGAGTTCATCTCTCGGGAGCATGCCAAG 1506  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCTCGCTGCGAGGCTGAGCTGGAAGATGAGCGTGGCGACTGGCTTGGCTGGCGAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCAGGGGTGGCTGTGTTCCGCGCGCAGAGCACCGCTCTGCGTGGAGGATCTCTGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1627 AAGTTCCTGCACTGGCTGATGATGTGTACGTGCTGCTGAGCTGCTCAGGTCTTTCTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTCAAGGAGACCAAGTTCCTCAAGAACACAGGCTCTTTTCTACCGGAAGAGTGTGGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTCGAAGCATTTGAATCAGACAGCACTTGAAGAGGTGTCAGTCTCGGAGCTGTGCG 1806  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAACAGAGAGTCAAGGAGCATCGGAAGCAGGCGCGCGCTCTGCTGACGCTCCAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1867 TTATCCCAAGCTCAAGGCTGCGGCTGCGCGATGTGAACATGAGTACGTCGTGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAAGCTTCGCGAGAGAAAGAGGCGGAGCGCTCTCACTCGAGGGTGAAGGCATGTTTC 1986  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 AGCGTGTCTCAACTACAGCGGCGCGCGCGCGCGCGCTCTCTGGGCGCTCTGTGCGGCG 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCAAGAGGCTTGGCGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CCGCTCTGAGCTGACTTGTCAAGGTGATGTGAGCGGCGCGCTCTCTGGGCGCTCTGTGCGG 2166  
713 GlyLysLeuValGlnValAlaLeuMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACAGGCTCAGGAGGTCTACCGCATCATCAAA-----CCCCAGAACACGATCTGGCGTG 2223  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2224 CGTCGCTATGCGCTGTCTCAGAAGCGCGCCATGGCACGCTCCGCAAGGCTTCAAGAGC 2283  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CAGCTCTCTACTTGAACAGCTCCAGCGCGTACATGCGACAGATTCGTGGCTCACCTCGAG

QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
DB 2344 GAGACC-----AGCCCGCTGAGGATGCGGTCTGTATCAGACGAGCTCTCCCTGAAT 2397  
QY 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
DB 2398 GAGGCGAGCAGTGGCTCTTCAGAGTCTTCCTACGCTTCATGTGCCACACGCGGTGGC 2457  
QY 813 IleGlyAspArgCysTyrThrGlnCysGlnIleProGlnIleProGlnIleSerSerThr 832  
DB 2458 ATCAGGGCAAGTCTCTACGTCAGTCCAGTGGCGGATCCCGAGGCTCCATCTCTCCACG 2517  
QY 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
DB 2518 CTGCTCTGAGCGCTGTGTACGGGACATGGAGAACAGCTGTGTGGGGGATTCGGCGG 2577  
QY 853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
DB 2578 GACGGCTGCTCTCTGGTGTGGTGAATTTCTGTGTGACACCTCACCTCACCCAC 2637  
QY 873 AlaTyrThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
DB 2638 GCGAAACCTTCTCTCAGACCTGCTCGAGGTCTCTGAGTATGCTCGGTGGTGAAC 2697  
QY 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
DB 2698 TTGCGAAGACAGTGGTGAATCTCTCTGTAGAAGACGAGCGCTGGGTGGACGGCTTTT 2757  
QY 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
DB 2758 GTTCAGATCGCGGCCACGGCTATTCTCCCTGTGGTGGCGCTGTGTGGATACCGGACC 2817  
QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
DB 2818 CTGGAGGTGCGAGGCACTACTCCAGCTATGCCGAGCTCCATCAGAGCCACTCTCACC 2877  
QY 953 PheGlnSerValPheLysAlaGlyThrMetArgAsnLysLeuLeuSerValLeuArg 972  
DB 2878 TTCAACCGCGCTCAAGGCTGGGAGAACATGCTGCGCAACTCTTTTGGGGTCTTGGCG 2937  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
DB 2938 CTGAAGTGTACAGCTGTTCTGTGATTTGAGGTGACAGCTCCAGCGGTGTGACCC 2997  
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
DB 2998 AACATCTACAAGATCTCTCTGCTGAGCGGTACAGGTTTCAAGCATGTGTGTGAGCTC 3057  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
DB 3058 CCATTTCTATCAGCAAGTTTGAAGAACCCACATTTTCTCTGCGCGTCACTCTTGACAGC 3117  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
DB 3118 GCCTCTCTCTGCTACTCTCATCTGAAAGCCAGAACCCAGGATGTGCTGGGGGCCAAG 3177  
QY 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
DB 3178 GCGCGCGCGCGCTCTGCTGCGCGCGCTGCGAGTGGCTGTGCGACCAAGCATTCCTG 3237  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
DB 3238 CTCAGCTGACTCGACCCGCTGCTACCTACGTGCGCATCTCTGGGCTCCTCAGACAGCC 3297  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
DB 3298 CAGACGAGCTGAGTGGGAAGCTCCCGGGAGCAGCTGACTGCTGCTGGAGCGCGCAGCC 3357  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
DB 3358 AACCGGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3396



```

433 ThrSerProHisLeuMetAspLeuArgLeuHisSerSerProTrrpGlnValTyr 452
1327 ACAGACCCCGTGGCTGGTGGAGCTGCTCCGCCAGACAGACAGCCCTTGGCAGGTGTAC 1386
453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTrrpGlyThrArgHis 472
1387 GGCTTCGTGGGGCTGCTGGCCGGCTGGTGGCCCGCCAGGCTCTGGGGCTCCAGGCAC 1446
473 AsnGluArgArgPhePheLeuLeuLeuLeuLeuPheLeuLeuLeuLeuLeuLeuLeu 492
1447 AACGAACCGCGCTTCCTCAGGAACACCAAGAGTTCATCTCCCTGGGGAAGCATCCCAAG 1506
493 LeuSerLeuGlnGluLeuMetTrrpLysMetLysValGluAspCysHisTrrpLeuArgSer 512
1507 CTCTCGCTGACAGAGCTGACGTGAAGATGAGCGTCCGGGATGCGCTGGCTGGCGGAGG 1566
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgGluArgLeuAla 532
1567 AGCCACAGGGTGGCTGTGTCGGCGCCAGAGCACCGTCTCGCTGAGGAGATCTCGGCC 1626
533 ThrPheLeuPheTrrpLeuMetAspTrrpValValGluLeuLeuArgSerPhePheTyr 552
1627 AAGTTCCTGCACTGGCTGATGATGTGCTGCTGAGCTGCTCAGGCTCTCTTTAT 1686
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrrpSer 572
1687 GTACGGAGACACAGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGATGCTGGAGC 1746
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1747 AAGTTCGAAAGCATTTGAATCAGACAGCATTTGAAGAGGGTGCAGCTGGCGGAGTGTG 1806
593 GlnGluGluValArgHisGlnAspTrrpLeuAlaMetProIleCysArgLeuArg 612
1807 GAACAGAGTTCAGGACATCGGAAGCCAGCCCGCTGCTGAGCTCCAGATCTCCGC 1866
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1867 TTCATCCCAAGCTGACGGGCTCGCGCCGATTTGTAACATGACATCTGCTGGGGAGCC 1926
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1927 AGAAGTTCGACAGAAAGAGGGCCGACGCTCTCCTCGAGGGTGAGGCATCTGTTTC 1986
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
1987 AGCTGTCAACTACGAGCGCGCGCGCCGCTCTCTGGCGCCCTCTGCTGGGCG 2046
673 MetAsnAspIleTyrArgThrTrrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2047 CTGACCATATCCACAGGGCTGGCGCACCTTCTGCTGCTGGTGGGGCCCGAGACCGG 2106
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712
2107 CCGCTGAGCTGACTTTGTCAAGTGTGATGTGACGGGGCGGTACGACACCATCCCCAG 2166
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2167 CACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGAACACGTACTCGCTG 2223
733 ArgGlnTrrpAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2224 CGTGGTATGCGGTGTCGAAGGGCCCGCCATGGCACGTCGCGAGGCCCTTCAGAGC 2283
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2284 CACGTCTCTACCTTGACAGACCTCCAGCCCGTACATGCGACAGTTCGTGGCTCACCTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2344 GAGACC-----AGCCCGCTGAGGATGTCGTCTCATCGACAGAGCTCTCCCTGAAT 2397
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812

```

```

2398 GAGGCAGAGCTGGGCTCTTTCAGCGTCTTCTACGCTTCATGTCGCCACCGCGGTGGC 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2458 ATCAGGGGCAAGTCTCAGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCACG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2518 CTGCTCTGACGCTGTGCTACGGCAGATGAGAACAGCTGTTTGGGGATTCGGCGG 2577
853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2578 GACGGGCTCTCTCGGTTTGGTGTGATTTCTTGTGTGTGACACCTCACCTCACCCAC 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2638 GCGAAACCTTCTCAGGACCTGTGTCGAGGTGCTCTGAGTATGCTGCTGGTGGTGAAC 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2698 TTGCGGAAGACAGTGTGNACTTCTCTGTAGAACGAGGCGCTGGGTGGCAGCGCTTTT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTrrpCysGlyLeuLeuLeuAspThrGlnThr 932
2758 GTTCAGATCGCGGCCAGCGCTATTCCTCTGGTGGCGCTGCTGCTGATACCGGAGC 2817
933 LeuGluValPheCysAspTyrSerGlyTrrpAlaGlnThrSerIleLysThrSerLeuThr 952
2818 CTGAGGTGCAGGCGACTTCTCCAGCTATGCCCGGACCTCATCAGAGCCAGTCTCACC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2878 TTCACCGCGGCTTCAAGGCTGGAGGAACATGCTGCGCAACTCTTTGGGCTCTTGGG 2937
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2938 CTGAAGTGTCAAGCGCTGTTCTGTGATTTGAGGTGAACAGCTCCAGACGGTGTGCACC 2997
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
2998 AACATCTACAAGATCTCTCTGCTGCGAGCGTACAGTTTCACGCATGTGTGTCAGGCTC 3057
1013 ProPheAspGluArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3058 CCATTTTCATCAGCAAGTTTGAAGAAGCCCACTTTTCTGCGGCTCATCTCTGCACAG 3117
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3118 GCCTCTCTGTGTTACTTCCATCTGAAAGCCAGAACGAGGATGTCGCTGGGGGCCAAG 3177
1053 GlySer-----PheProProGluAlaAlaHisTrrpLeuCysTyrGlnAlaPheLeu 1059
3178 GGGCGCGCGGCTCTCTGCTCCGAGCGCGTGCAGTGGCTGTGTCACCAAGCATTCCTG 3237
1070 LeuLysLeuAlaAlaHisSerValIleTrrpLysCysLeuLeuGlyProLeuArgThrAla 1089
3238 CTCAAGCTGATCGACACCGTGTACCTACGTGCGACCTCTCTGGGTCTACTCAGACAGCC 3297
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
3298 CAGACCGAGCTGAGTCGGAAGCTCCCGGGGAGCAGCGTGACTGCCCTGGAGCGCGCAGCC 3357
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3358 AACCCGCGCATGCCCTCAGACTTCAAGACCATCTCTGGAC 3396

```

RESULT 5

ID ABZ35720 standard; DNA; 3399 BP.

XX ABZ35720;

AC ABZ35720;

XX



61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

191	DB	GACCACGGCGCCCGCCCGCCCTCTTCGCCGAGGTGTCTCTGACGAGAGCTG	240
81	QY	ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
241	DB	GTGCGCCGAGTGTGCAGAGGTGTGCAGGCGCGCGCGAAGACGTCGTGGCTTCGGC	300
101	QY	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetaAlaPheThrSerSerValArg	120
301	DB	TTTCGGTGTGTGACGGGCGCGGGGGCCCCCGAGGCTTTCACCCACGCGTGGC	360
121	QY	SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu	140
361	DB	AGCTACTGTGCCAACACGCTGACCGACGACATGCGGGGACGCGGGCGTGGGGCTGCTG	420
141	QY	LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu	160
421	DB	CTGCGCGCGTGGCGGACGAGCTGTGTTCACCTGCTGGCACGCTCGCGCTCTTTGTG	480
161	QY	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla	180
481	DB	CTGTGGGTCCACGCTCGCGCTACCAAGTGTGCGGGCGCGCGCTGTACCAAGCTCGCGCT	540
181	QY	ThrThrAspIleTTPProSerValSerAlaSerTyrArgProThrArgProValGlyArg	200
541	DB	GCACCTCAGGCGCGCGCCCGCCACACGCTAGT---GGACCCCGAAGCGCTCTGGGA---	594
201	QY	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro	220
595	DB	-----TCCGACCGCGCTCGAACCATAGCTGAGGAGGCGCGG	633
221	QY	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer	240
634	DB	GTCCCGCTGGCGCTGCCAGCCCGGGTGGCAGGAGCGCGGGCGAGTGCACGCGAAGT	693
241	QY	ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro---	259
694	DB	CTGCGGTGTGCCAAGAGCCGACGCGTGGCGTCCCTGACGCGGACGCGACGCCCGT	753
260	QY	-----HisArgGlnValLeuProThrProSerGlyLysSerTrp---	272
754	DB	GCGCAGGGGTCTTGGGCCCCACCGGGCAGCAGCGTGGACGAGTACCGTGTCTTGT	813
273	QY	ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys	292
814	DB	GTGGTGTCACTGCGAGA-----CCGCGCGAAGAAGCACCTCTTTGGAG	858
293	QY	GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer	311
859	DB	GGTGGCTCTTGGCAGCGCCACTCCACCCACTCGTGGCGCGCCAGCACACGCGCGC	918
312	QY	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---Phe	330
919	DB	CCCCATCCATCCGCGGCCACACGTCCTCGGACACGCTTGTCCCCCGGTGTAGCC	978
331	QY	GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnArgLeuAsnProSerPhe	350
979	DB	GAGACCAAGCACTTCTACTCTCTCAGGCGAC---AAGGACAGCTCGCGCCCTCCTC	1035
351	QY	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe	370
1036	DB	CTACTCAGCTCTTAGGCCCCAGCTGACTGCGCTCGGAGCTCTGTGAGACCATCTTT	1095
371	QY	LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg	390
1096	DB	CTGGGTTCAGGCCCTGGATGCCAGGACTCCCGCAGGTGTGGCGCGCTGCCCGCGC	1155
391	QY	TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr	410
1156	DB	TACTGCAAAATCGGCCCCCTTTCTGAGACTCTTGGGAACCAACGCGAGGTGCCCTAC	1215
411	QY	ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla	430
1216	DB	GGGGTCTCTCAAGACCACTGCGCGCTGAGCTGCG-----GTACCCCGACA	1266

431 Leu-----Asn 432  
1267 GCCGGTGTGTGCCCCGGGAGAACCCAGGGCTGTGTGGCGGCCCCCGAGGAGGAGGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1327 ACAGACCCCGTGGCTGGTGTGAGCTGCTCCGCGCAGCACAGACGCCCTGGCAGGTGTAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLeuValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GCGTTCGTGGCGGCTGCTCCGCGCGCTGGTGGCCCGCCAGGCCCTCTGGGGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheLeuAsnLeuLysPheIleSerLeuGlyLysTyrGlyLys 492  
1447 AACGAAACCGCTTCCTCAGCAACCAACCAAGATTCATCTCCCTGGGGAAGCATGCCAAG 1506  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGGCGGACTGCGCTTGGCTGGCGCAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCAGGGGTGTGCTGTGTCCGGCCGCGAGACACCGTCTGGTGAGGAGATCTCGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1627 AAGTTCCTGCACTGGCTGATGAGTGTGTACGTGTGTGCGAGCTGCTCAGGTCTTCTTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTCAGGAGACACAGCTTCAAAAGAACAGGCTCTTTTCTACCGGAAGATGTCGTGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTGCAAGCAATGGATCAGACAGCACTTGAAGAGGGTGCAGCTGGCGGAGCTGTGCG 1806  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAAGCAGAGGTGAGCAGCATCGGGAAGCCAGCGCCCGCTGTGAGCTGCAGCTCCAGCTCCG 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1867 TTCATCCCAAGCTGACGGCTGCGGCCGATTTGAACATGACATGACATGCTGCGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAACGTTCCGCGAGAGAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTTC 1986  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 AGCGTGTCACTACGACGGCGGGCGGGCGCCCGGCTCTCTGGGGCGCTCTGTGCTGGGC 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCCAGGGCTGCGGCACCTCTGCTGTGTGCGGTGCGGGCCAGGACCCG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CCGCTGTAGCTGTATTTGTCAAGTGTGATGTGAGGGCGCGGTACGACACCATCCCCCAG 2166  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACGGCTCAGGAGGTTCATCGCAGCATCATCAAA---CCCCAGAACAGTACTTGGGTG 2223  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2224 CGTGGTATGCCGTGCTCCAGAGGAGCCCGCCATGGGCAAGCTCCGCAAGGCTTCAAGAGC 2283  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CAGCTCTACCTTGACAGACCTCCAGCGCTGATGCGACAGTTCGTGGCTCACCTGCGAG 2343

773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluInSerIleSerMetAsn 792  
2344 GAGACC-----AGCCCGGTGAGGATGCCGTGTCATCGAGCAGAGCTCCTCCTGTAAT 2397  
793 GluSerSerSerSerLeuPhePhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAGCCAGCAGTGGCTCTTCGACGCTTCTCTACGCTTCATGTGCCACACGCGCTGGC 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCAGGGGCAAGTCTTACGTCAGTCCAGGCCAGGGATCCCGAGGGTCTCATCTCTCCAG 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 CTGCTCTGAGCCCTGCTACGCGCAGCATGGAGAAACAAGCTGTTTGGCGGGATTCGGGG 2577  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2578 GACGGGCTGCTCTGCGTGTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCAC 2637  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 GCGAAACCTTCTCAGGACCCCTGGTCCGAGGTGCTCCCTAGTATGGCTGCGTGGTGAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGGAAGACAGTGTGAACCTTCCCTGTAGAGACGAGGCCCTGGGTGGCAGCGCTTTT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATGGCGGCCCAACCGGCTATTCCTCTGGTGGGCTGCTGCTGATACCCGGACC 2817  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGAGGTGACAGCAGCACTACTCCAGCTATGCCGAGACCTCCATCAGAGCCAGTCTCACC 2877  
953 PheGlnSerValPheLysValGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGCGGCTTCAAGGCTGGAGGAACATGCTGCGCAACTCTTTGGGGTCTTGGCG 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 CTGAAGTGTACAGCCTGTTCTTCGATTTGCAAGTGAACAGCCTCCAGACGGTGTGCACC 2997  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATCTACAGATCTCTCTGCTGAGGGGTACAGGTTTTCACGATGTGTGCTGAGCTC 3057  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTTTCATCAGCAAGTTTGAAGAAGACCCACATTTTCTGCGCGTCTATCTGTACACG 3117  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCCCTCTGCTACTCCATCTCTGAAGCCAAAGACGAGGGATGTCGCTGGGGCCCAAG 3177  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3178 GCGCGCCCGGCGCTCTGCGCTCCGAGGGCGGTGACGTGGCTGTGCCACCAAGCATTTCTTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 CTCAAGCTGATCGACACCGGTGTACGTAGTGCCTCTCTGGGGTCACTCAGGACAGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3298 CAGACGAGCTGAGTCCGAAGCTTCCCGGGGACGACGCTGACTGCCCTGGAGCCGCGAGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCCGCGACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3396

RESULT 6

:09963

ABX09963 standard; DNA; 3399 BP.

ABX09963;

23-JAN-2003 (first entry)

Human telomerase reverse transcriptase DNA fragment SEQ ID 28.

oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;  
prion; inhibition; human; ds.

**Homo sapiens.**

DE10100587-C1.

21-NOV-2002.

09-JAN-2001: 2001DE-1000587.

09-TAN-2001: 2001DE-1000587.

(RIBO-) RIBOPHARMA AG.

Kreutzer R, Limmer S, Rost S, Hadwiger P;

WPI: 2002-742209/81.

Inhibiting expression of target genes, e.g. oncogenes, in cells, by introduction of complementary double-stranded oligoribonucleotide, after treating the cell with interferon -

Disclosure: Page 30-31; 98pp; German.

This invention describes a novel method for inhibiting expression of a target gene by introducing into the cell that contains the target gene at least one oligonucleotide (dsRNAi) that has a double-stranded (ds) structure of not more than 49 consecutive nucleotides (nt), where at least a segment of one strand of the ds structure is complementary with the target gene and the cells are treated with interferon before introduction of dsRNAi. The method is used to inhibit expression of target genes, particularly oncogenes, cytokine genes, Id (not defined) protein genes, developmental or prion genes, or genes expressed in pathogenic organisms (particularly plasmodia) or in viruses or viroids (pathogenic in humans, animals or plants). Treating the cells with interferon greatly increases the extent to which dsRNA can inhibit expression of the target genes, and the effect is even greater when dsRNA are modified to increase their stability. ABX09336-ABX10075 represent gene fragments used to illustrate the method of the invention.

Sequence 3399 BP: 550 A: 1157 C: 1089 G: 603 T: 0 other:

ianment Scores:

Length:	9.5e-234	3399
Matches:	3505.00	719
Conservative:	72.94%	132
Mismatches:	62.34%	260
Indels:	59.40%	52
Gaps:	24	13

-09-042-460-2 (1-1122) x ABX09963 (1-3399)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
1 ATGCCGGCGCTCCCGCTGCCAGCCGTCCTCTCCCTGCGCAGCACATCCCGCAG 50  
21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
61 GTGCTCCCGCTGCCACGTCTGTGCGCGCCCTGGGCGCCCGCGGCTGGCGGTGTGCGAG 120  
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60  
121 CGCGGGGACCGCGCGCGCTTTCGCGCGCTGTGTGGCCAGTGCTGTGTGTGCGTGGCGCTG 180

411 ValArgLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1216 GGGGTGCTCTCAGACGACACTGCCCGGTGAGCTGG-----GTACCCAGCA 1266  
431 Leu-----Asn 432  
1267 GCGGTGTCTGTGCGCGGAGAAAGCCCGAGGCTCTGTGGCGGCGCCCGAGGAGGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1327 ACAGACCCCGTGGCTGTGTGAGCTGTCTCCGCCACGACAGACGCCCTTGGCAGGTATC 1386  
453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGTCTGTGCGGGCTGCTGCGCGGCTGTGTGCGCCAGGCTCTGGGGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheHisLeuLeuLysLysPheLeuLysLysTyrGlyLys 492  
1447 AACGAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCTCTGGGGAAGCATGCCAAG 1506  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTTCTGCTCAGAGAGCTGAGCTGGAGATGAGCGTGGGACTGGCTTGGCTGGCAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 AGCCAGGGGGTGGCTGTGTTCCGCGCCGACAGACACCGCTCTGCGTGAGGAGATCCTGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTrpValValGlnLeuLeuArgSerPhePheTyr 552  
1627 AGTTCTGCATGGCTGATGATGTGTGACGTCTGCTGAGCTGCTCAGTCTTTCTTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTCACGGACACCACTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AGTTCAAGCATTTGGAATCAGACAGACACTTGAAGAGGGTGCAGCTGCGGAGCTGCG 1806  
593 GlnGluValArgHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
1807 GAAGCAGAGGTGAGGAGCATCGGGAAGCAGCGCCGCTGCTGACGTCCAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProLysValAlaAsnMetSerTrpSerMetGlyThr 632  
1867 TTCATCCCAAGCCTGACGGGTGCGCGGATGTGAACATGGAATGAGTACGTCTGGGAGCC 1926  
633 ArgAlaLeuGlyValArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAACCTCCGACAGAAAGAGGGCCGAGCGCTCTCCTCGAGGGTGAAGCACTGTT 1986  
653 SerMetLeuAsnTrpGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 AGCGTCTCAACTACAGCGGGCGGCGCGCCCGCCCTCTCTGGCGGCTCTGTCTGGGC 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGAGCATATCCAGAGGCTTGGCGACCTTGTGCTGCGTGTGCGGGCCAGAGCCCG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CCGCTGTAGCTGTACTTGTCAAGGTGATGTGACGGCGCGGTACGACCACTCCCGAG 2166  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACAGCTCAGGAGGTATCCGCCAGCATCATCAAA-----CCCCAGAACACGATCTGCGTG 2223  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 GTTCGGTATGCCGTGCTCCAGAGGCGGCCCATGGGCACGTCGCCAAGGCCCTTCAAGAGC 2283  
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772

2284 CACGTCTCTACCTTGAAGACCTCCAGCGCTACATCGACAGTTCGTGGCTCACCTGCAG 2343  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2344 GAGACC-----AGCCCGCTGAGGATCGCGTCTCATCGACAGAGCTCTCTCCCTGAAT 2397  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAGCGCAGCAGTGGCTCTTCGACGCTTCTTACGCTTTCATGTGCCACACCGCGTGGC 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCAGGGCAAGTCTACGTCTCAGTGCAGGGATCCCGCAGGGCTCCATCTCTCCACG 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 CTGCTCTCGACCTGTCTACGGGACATGAGAACAGCTGTTTGGGGGATTCGGCGG 2577  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2578 GACGGGTGCTCTCTGCGTGGTGTGATGATTTCTTGTGTGTGACACCTCACCTCACCCAC 2637  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 GCGAAACCTTCTCCAGACCCCTGTCGAGGGTGTCCCTGAGTATGCTGCGTGGTGAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGGAAGACAGTGGTGAATCTCCCTGTAGAACAGAGCGCCCTGGGTGGCAGGCTTT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATCGCGGCCACCGCTTATTCCTGTGGTGGCGCTGTGTGGATACCCGAGC 2817  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGGAGGTGCAGAGGACTACTCCAGCTATGCCGACCTCCATCAGAGCCAGCTTCACC 2877  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGCGCTTCAAGGCTGGGAGGAACATCGTCGCAAACTCTTTGGGGTCTTGGCG 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuValAsnSerLeuGlnThrValCysIle 992  
2938 CTGAAGTGTACACGCTGTTTCTGATTTGAGGTGAGAGCCCTCCAGCGGTGTGCACC 2997  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATCTACAAGATCTCTCTGCTCAGCGGTACAGGTTTTCACGCATGTGTGTGCAGCTC 3057  
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTTCATCAGCAAGTTTGGAGAACCCCACTTTTCTCTGCGCGTCATCTCTGACAGC 3117  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCTTCTCTCTGCTACTCCATCTGAAAGCCAGACGAGGATGTGCTGGGGGCCAAG 3177  
1053 GlySer-----PheProProGluAlaAlaHisTrpLysCysLeuLeuGlyProLeuArgThrAla 1089  
3178 GCGCGCGCGCGCTCTGCGCTCCGAGGCGGTGCGAGTGGCTGTGCCACCAAGCATTCCTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 CTCAGCTGATCCAGACCCGCTGTCACCTAGTCCACTCTCTGGGGTCACTCAGAGCAGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3298 CAGACCGAGCTGAGTCCGAAGCTCCCGGGACGACGCTGACTGCTGAGCGCGGAGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122



1156 TACTGGCAATGCGGCCCTCTTTCTGAGCTGTTGGGAACAACGCGAGTCCCTTAC 1215  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAlaGlnGlnValThrAspAla 430  
1216 GGGTGTCTCTCAAGACCACTGCCGCTGAGCTGG-----GTCAACCCAGCA 1266  
431 Leu-----Asn 432  
1267 GCGGTGTCTGTGCGCGGGAAGCCCGAGGGCTCTGTGCGGCCCGCGAGGAGGAC 1326  
433 ThrSerProProHisLeuMetAspLeuLeuHisSerSerProTrpGlnValTyr 452  
1327 ACAGACCCCGCTGCGCTGCTGCGCGCGCTGCTGCGCGCGAGCCCGCTGGCAGGTGTAC 1386  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGTCTGTGCGGGCTGCTGCGCGCGCTGCTGCGCGCGCTGCTGCGCGCTCCAGGCAC 1446  
473 AsnGluArgArgPhePheLeuAsnLeuLysPheLeuSerLeuGlyLysTyrGlyLys 492  
1447 AACGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCCCTGGGGAAGCATGCCAAG 1506  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTCCTGCTGCGAGGCTGACGTGGAAGATGAGCGTGGGACTGCGCTTGGCGCGCAGG 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLysLeuAla 532  
1567 AGCCAGGCGGTGGCTGTTTCCGCGCGCAGAGCACCGTCTGCGTGGAGAGATCTCGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1627 AAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGCTGCGTGTGCTGCTGCTTTTAT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTACGCGAGACACGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAGTTGCAAGCAATTGGAATCAGACAGCACTTGAAGAGGCTGCGAGTGGCGAGCTGCG 1806  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAGCAGAGGTCAGACGATCGGGAAGCCAGCCCGCTCTGTGAGCTCCAGACTCCGC 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1867 TTCAATCCCAAGCTGACGGGCTGGGCGGATGTGAACATGGACTACGTCGTGGGAGCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 AGAACGTTCCGCGAGAAAGAGGCGGAGGCTCTCACTCGAGGGTGAAGGCACCTGTTTC 1986  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 ACGTGTCTCACTACGACGGGCGGCGCGCCCGCTCTGCGCGCTCTGTGCTGGGCG 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGGACGATATCCACAGGCGCTGGCGCACCTTCTGTGTGCTGTGGGCGCCAGGACCG 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CGCGCTGAGCTGTACTTGTCAAGTGGATGTACGGGCGCGTACGACACCATCCCGCCAG 2166  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGlnSerThrTyrCysIle 732  
2167 GACAGGCTCAGGAGGTCTACCGCAGCATCATCAA---CCCCAGAACAGTACTCGGTG 2223  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752

Db 2224 CGTCGTATCGCTGGTCCAGAGGCGCCCATGGGCACGTCGCCAAGGCCTTCAAGAGC 2283  
Qy 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2284 CAGCTCTCTACTTGACAGACCTCCAGCGGTATCATGCGACAGTTCGTGGCTCAGCTGAG 2343  
Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2344 GAGACC-----AGCCCGCTGAGGGATGCGTCTCATCGACAGAGTCTCTCCCTGAAT 2397  
Qy 793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
Db 2398 GAGCCAGCAGTGGCTCTCTGACGCTCTTCCACGCTTATGTCGCCACCCAGCCGTCGCGC 2457  
Qy 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
Db 2458 ATCAGGGGCAAGTCTACGTCCAGTCCAGTCCAGGGATCCCGCAGGGCTCCATCTCTCCAG 2517  
Qy 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
Db 2518 CTGCTCTGACGCTGTGCTACGCGCAGCATGAGAACAGAGCTGTTTGGCGGGATTCGCGCG 2577  
Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
Db 2578 GACGGCTGCTCTGCGTGTGGTGGATGATTTCTTGTGTGACACCTCACCCTCACCAC 2637  
Qy 873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
Db 2638 GCGAAACCTTCTCAGGACCTTGGTCCGAGGTGTCCTGAGTATGGCTGCTGCTGTAAC 2697  
Qy 893 LeuGlnLysThrValAlaAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
Db 2698 TTGCGAAGACAGTGTGTAACTTCCCTGTAGAAGACAGGCGCTGGGTGGCAGCGCTTTT 2757  
Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2758 GTTCAGATGCGCGCCACCGCCTATTCCCTCGTGGTGGGCTGCTGCTGGATACCCGACC 2817  
Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2818 CTGGAGGTGACAGCAGCACTACTCAGCTATGCGCGAGCCTCCATCAGACGCACTCTCACC 2877  
Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
Db 2878 TTCAACCGCGCTTCAAGGCTGGAGAGAACATGCTGCGCAAACTCTTTGGGGTCTTGCG 2937  
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2938 CTGAAGTGTCAACGCTGTTTCTGGATTTGCAAGTGAACAGCCTCCAGACGCTGTGCACC 2997  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 2998 AACATCTACAGATCTCTGCTGTCAGGCGTACAGGTTTTCACGCAATGTGTGTGCGACTC 3057  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3058 CCATTTCATCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGCTCATCTCTGACAGC 3117  
Qy 1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaser 1052  
Db 3118 GCCTCTCTGCTACTCTCATCTCTGAAGCCAGAACAGCAGGATGTGCTGGGCGCCAG 3177  
Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
Db 3178 GCGCCGCGCGCGCTCTGCGCTCCGAGGCGCGTGCAGTGGCTGTGCCCAAGCATTCCTG 3237  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3238 CTCAAGCTGACTCGACACCGTGTCACTAGTGCACCTCTCTGGGGTCACTCAGGACAGCC 3297  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3298 CAGACGCACTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCTGCTGCGCTGGAGCGCCAGCC 3357

1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
:::|||||:::|||||:::|||||:::|||||  
3358 AACCGGCACCTGCCTCAGACTTCAAGACCATCTGGAC 3396

SULT 8  
L91685

ABL91685 standard: DNA: 3399 BP.

ABL91685:

28-MAY-2002 (first entry)

Human polynucleotide SEO ID NO 28.

Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen; Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide; cytostatic; viricide; protozoicide; antibacterial; ds.

**Homo sapiens.**

DE10100586-C1.

11-APR-2002.

09-JAN-2001: 2001DE-1000586.

09-JAN-2001: 2001DE-1000586.

(RIBO-) RIBOPHARMA AG.

Kreutzer R. Limmer S. Rost S. Hadwiger P:

WPI: 2002-270454/32.

Inhibiting gene expression in cells, useful for e.g. treating tumors, by introducing double-stranded complementary oligoRNA having unpaired terminal bases -

Claim 13: Page 27-28: 104pp: German:

The invention relates to a method for inhibiting expression of a target gene (ABL91658-ABL91797) in a cell by introducing at least one oligoribonucleotide that has a double-stranded structure consisting of at most 49 sequential nucleotide pairs, with at least part of one strand complementary with the target gene and has at least one end a single-stranded segment of 1-4 nt. The method provides oligoribonucleotides for antisense inhibition of gene expression useful e.g. for treating tumours but the oligoribonucleotides may also be directed against genes present in pathogens (e.g. Plasmodium or viruses/viroids, pathogenic to humans, animals or plants) or against cytokine, *lfd*, developmental or *pion* genes. The method provides more effective inhibition of gene expression than use of known oligonucleotides, probably because the unpaired overhang increases stability and thus intracellular concentration.

Sequence 3399 BP: 550 A; 1157 C; 1089 G; 603 T; 0 other;

ignment Scores:	
ed. No.:	9.56-234
ore:	3505.00
Percent Similarity:	72.94%
st Local Similarity:	62.36%
ery Match:	59.40%
};	24
	Length:
	Matches: 3399
	Mismatches: 719
	Conservative: 122
	Mismatches: 260
	Indels: 52
	Gaps: 13

;-09-042-460-2 (1-1122) X ABL91685 (1-3399)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
1 ATGCCGCGGCTCCCGCTGCCGAGCGTGGCTCTGCGCAGCCACTACCGCCGAG 60  
21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluValArgArgLeuValGln 40

21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40

	61	GTGCTGCGCTGCGCACAGTTGCTGTCGCGCCCTTGGGGCCCCAGGGCGCTGCGCGCTGGTGCGAG	120
Db			
Qy	41	ProGlyAspProLysIleYrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp	60
Db			
Qy	121	CGCGGGGACCCGGCGGCTTTCCGGCGCTGTGTGGCCAGTGCCTGTTGCTGCGCTGCCCTGG	180
Db			
Qy	61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
Db			
Qy	181	GACGACGGCGCCCGCCCGCCCTCTTCCGCCAGGTGCTCTGCTGAGAGAGCTG	240
Db			
Qy	81	ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
Db			
Qy	241	GTGCCCCAGTGTCTGCAGAGCTGTGCGAGCGCGCGCGCGGCGGAAGAACTGCTGCGCTTCGCG	300
Db			
Qy	101	PheGluLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120
Db			
Qy	301	TTTCGCTGTGTGACGCGGGCCCGCGGGGCGCCCGCCAGGCTTCACACACAGCGTGGCG	360
Db			
Qy	121	SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu	140
Db			
Qy	361	AGCTACTTGCCCAACAGGTGACCGACGACCTGCGGGGGAGCGGGGCGTGGGGGCTGGT	420
Db			
Qy	141	LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu	160
Db			
Qy	421	CTGCGCGCGTGGCGACGACGCTGCTTACCTGCTGGCACGCTGGCGCTCTCTTTGTG	480
Db			
Qy	161	LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla	180
Db			
Qy	481	CTGTGGCTCCAGCTCGGCTACCAAGGTGTGCGGGCCGCTGTATCAGCTCGGCGCT	540
Db			
Qy	181	ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg	200
Db			
Qy	541	GCCACTCAGGCGCGGCCCGCCACACGCTAGT--GGACCCCGAAGCGCTGTGGCA---	594
Db			
Qy	201	AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro	220
Db			
Qy	595	-----TCGCAACGGCGCTGGAACCATAGTCGTGAGGAGCCCGGG	633
Db			
Qy	221	LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer	240
Db			
Qy	634	GTCCCCCTGGCGCTGCAGCCCGGGTGGGAGGAGCGCGGGGCGAGTGCCACGCCAAGT	693
Db			
Qy	241	ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro---	259
Db			
Qy	694	CTGCGCTGTGCCAAGAGGCCAGCGCTGCGCTGCGCCCTGAGCCGAGCGGACGCCCGCT	753
Db			
Qy	260	-----HisArgGlnValLeuProThrProSerGlyLysSerTrp---	272
Db			
Qy	754	GGCGAGGGCTCTGGGCCCAACCGGCGCAGGACCGCTGGACCGAGTGACCGTGGTTCTGT	813
Db			
Qy	273	ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys	292
Db			
Qy	814	GTGGGTGCACCTGCCAGA-----CCGCGCGAAGACCCACCTCTTTGGAG	858
Db			
Qy	293	GlyLysValSerAspLeuSerLeuSer--GlySerValCysCysLysHisLysProSer	311
Db			
Qy	859	GCTGGCTCTCTGGCACCGCGCCACTCCCAACCATCCGTGGCGCCGACGACCCACCGGGC	918
Db			
Qy	312	SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---Phele	330
Db			
Qy	919	CCCCCATCCACATCGCGGCCACACAGTCCCTGGGACACGCTTGTCCCGCGGTGACGCC	978
Db			
Qy	331	GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe	350
Db			
Qy	979	GAGACCAAGCACTTCTCTACTCTCTCAGGCGAC---AAGAGCAGCTGGCGCCCTCTTC	1033
Db			
Qy	351	LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe	370
Db			
Qy	1036	CTACTCAGCTCTCTGAGGCCCGAGCTGACTGGGCTCGAGGCTCTGGAGACCATCTTT	1095
Db			
Qy	371	LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg	390
Db			



1096 CTGGGTCCAGGCCCTGGATGCCAGGAGCTCCCGCAGAGTTGCCCGCCCTGCCCCAGCGC 1155  
 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 1156 TACTGGCAAAATGGCGCCCTGTTTCTGGAGCTCTCTGGGAACCAAGCGCAGTGCCTTAC 1215  
 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 1216 GGGGTGCTCTCAGACGCACTGCCCGCTCGAGCTGCG-----GTACCCACGCA 1266  
 431 Leu-----Asn 432  
 1267 GCCGGTGTCTGTGCCCGGAGAAAGCCCGAGGCTCTGTGGCGGCCCGCAGGAGGAGAC 1326  
 433 ThrSerProHisIleuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 1327 ACAGACCCCGTGGTGGTGGAGCTCTCCGCGCAGCAGAGCCCTGGAGAGTGTAC 1386  
 453 GlyPheLeuArgAlaCysLeuCysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 1387 GCGTTCGTGGGGCTCTGCTGGCGCGCTGCTGCCCGCCCTCTGGGCTCCAGGCAC 1446  
 473 AsnGluArgArgPhePheLeuLeuGlyPheIleSerLeuGlyTyrGlyLys 492  
 1447 AACGAACCGCGCTTCTCCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1506  
 493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 1507 CTCTCGTGCAGGAGCTGACGTGAAGATGAGCGTGGGGAGTGGCTGGTGGCAGG 1566  
 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
 1567 AGCCACAGGGTGGCTGTGTCCGGCGCAGAGCAGCGTCTGGCTGAGGAGATCTGGGCC 1626  
 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 1627 AAGTTCCTGCACTGGCTGATGAGTGTGTAGCTGTGAGCTGTCTAGGCTCTTCTTTAT 1686  
 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
 1687 GTCACGGAGACCATGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1746  
 573 LysLeuGlnSerIleGlyValArgGlnHisGluArgValArgValArgLeuArgGlnLeuSer 592  
 1747 AAGTGTCAAAAGCATTGAATCAGACACACATCTGAAGAGGGTGCAGCTGGCGGAGCTGTCG 1806  
 593 GlnGluValValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
 1807 GAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCCCTGCTGAGCTCCAGACTCCGC 1866  
 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
 1867 TTCATCCCAAGCCTGACGGCTCGCGCCGATGTGAACATGGACTACGTCGTGGGAGCC 1926  
 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
 1927 AGACCTTCCCAAGAAAGAGAGGCCCGAGCTCTCACCTCGAGGGTGAAGCAGCATGTTTC 1986  
 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
 1987 AGCGTGTCAACTACGAGCGGCGCGCGCGCCGCTCTCTGGCGCCCTCTGTGTGGGC 2046  
 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
 2047 CTGGACGATATCCACAGGGCTGGCGCACCTTGTGTGGTGTGGGGCCCGCAGGACCG 2106  
 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValArgAspAlaIleProGln 712  
 2107 CCGCTGAGCTGTACTTGTCAAGTGTGTGACGGCGCGTACGACCATCCCTCCCCAG 2166  
 713 GlyLysLeuValGlnValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
 2167 GACAGGCTCACGGAGGTTCATCGCGAGCATCATCAA-----CCCCAGAACAACGTACTGCGTGT 2223

Qy 733 ArgGlyTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
 Db 2224 CGTCGGTATCGTGGTCCAGAGGCCCGCATGGGCACGTCGCGAAGCGCTTCAAGAGC 2283  
 Qy 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
 Db 2284 CAGCTCTCTACCTTGACAGACCTCCAGCGGTACATCGCACAGTTCGTGGCTCACCTGAC 2343  
 Qy 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
 Db 2344 GAGACC-----AGCCCGCTGAGGATCGCTGTCATCGACGAGAGCTCTCTCCCTGAAT 2397  
 Qy 793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
 Db 2398 GAGCCAGCAGTGGCTCTTTCAGCGCTTCTTACCGTTCATGTGCCACACCGCGTGGCG 2457  
 Qy 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
 Db 2458 ATCAGGGGCAAGTCTCTACGTCCAGTGCAGGGGATCCCGCAGGGTCCATCTCTCCACG 2517  
 Qy 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
 Db 2518 CTGCTCTGAGCCTGTGCTACCGCGCAGCATGAGAGAAAGCTGTTCGCGGGATTCGGCGG 2577  
 Qy 853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
 Db 2578 GACGGGCTGCTCTCGCTTGTGGTGGATGATTTCTTGTGGTGACACCTCACCTCACCCAC 2637  
 Qy 873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
 Db 2638 GCGAAACCTTCTCTCAGGACCTGTGTCGAGGTGCTCCCTGAGTATGGCTGGGTGGTGAAC 2697  
 Qy 893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
 Db 2698 TTGCGGAAGACAGTGGTGAACCTTCCCTGTGTGGTGGCGCTGGTGGCGCGCTTTT 2757  
 Qy 913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
 Db 2758 GTTCAGATGCGCGCCACCGCCTATTCCCTGTGTGGTGGCGCTGTCTGTGGATACCGCGAC 2817  
 Qy 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
 Db 2818 CTGAGGTGACAGCGACACTCTCCAGCTATGCCCGAGCTCCATCAGAGCCAGTCTCAC 2877  
 Qy 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
 Db 2878 TTCACCGCGGCTTCAAGGCTGGAGAACATGCTGCGCAAACTCTTTGGGCTCTTGGCG 2937  
 Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
 Db 2938 CTGAAGTGTACAGCCTGTTCCTGGATTTCAGGTGAACAGCTCCAGCGGTGTGAC 2997  
 Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
 Db 2998 AACATCTACAAGATCTCTGCTGTCAGCGGTACAGGTTTCAGCGCATGTGTGTCGAGCTC 3057  
 Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
 Db 3058 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTCTGCGCGTCTATCTGTGACACG 3117  
 Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
 Db 3118 GCTCTCTCTGTCTACTCTCATCTCTGAAGCAAGACGAGGATGTCGTGGGGGCCAAG 3177  
 Qy 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
 Db 3178 GGGCGCGCGCGCCCTCTGCGCTCCGAGGCGGTGACGTGGCTGTGCGCACCAAGCATTCCTG 3237  
 Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
 Db 3238 CTCAAGCTGACTCGACCGCTGTCACTACGTGGCCACTCTGGGGTCACTCAGGACAGCC 3297

1090 GlnLysLeuLeuCysArgGlyLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
 3298 CAGACGAGCTGAGTCGGAAGCTCCCGGACGAGCGTACTGCTGGAGCGCGAGCC 3357  
 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
 3358 AACCCGGCACTGCCCTCAGACTTCAGACCAATCTGGAC 3396

MULT 9  
 727876

AAV27876 standard; cDNA; 3798 BP.

AAV27876;

25-MAR-2003 (updated)  
 12-OCT-1998 (first entry)

Human telomerase protein 2 (TP2) full-length cDNA.

TP2; human; telomerase protein 2; cancer; AIDS; ageing; therapy;  
 SS.

Homo sapiens.

Key Location/Qualifiers  
 CDS 2..3466  
 /tag= a

WO9821343-A1.

22-MAY-1998.

13-NOV-1997; 97WO-US21248.

15-NOV-1996; 96US-0751189.

11-JUN-1997; 97US-0873039.

16-OCT-1997; 97US-0951733.

(AMGE-) AMGEN CANADA INC.

(AMGE-) AMGEN INC.

Harrington LA, Robinson MO;

WPI; 1998-297946/26.

P-PSDB; AAV27876.

New nucleic acid encoding human telomerase protein-2 - used for  
 regulating telomerase activity, e.g. for treating cancer or acquired  
 immune deficiency syndrome

Claim 1c; Fig 9; 150pp; English.

This full-length cDNA clone codes for a human telomerase protein  
 2 (TP2, see AAW61350), a novel protein of the telomerase complex.  
 The sequence was deduced from overlapping partial clones #32 (see  
 AAV27876) and TP2-15 (see AAV27876), which were obtained from a human  
 colon tumour cell line LIM1863 cDNA. Expressing TP2 genes in a  
 cell is used to increase telomerase activity and thus proliferation  
 for treatment of e.g. HIV infection, AIDS and ageing disorders,  
 while expressing an inactive mutant of TP2 (or molecule antisense  
 to the gene) is used to decrease telomerase activity, e.g. for  
 treatment of cancer. TP2 polypeptides can also be used to screen  
 for agents that inhibit TP2 activity or its binding to TRP1 (see  
 AAW61347) or telomerase RNA, potentially useful therapeutically,  
 also to raise specific antibodies useful in immunoassays and  
 therapeutically as inhibitors. Nucleic acid fragments are used  
 as diagnostic probes for detecting/quantifying TP2 DNA. Also  
 contemplated are transgenic animals in which the TP2 gene has been  
 inactivated or is overexpressed. TP2 polypeptides are administered  
 i.v., s.c. or orally, or they are delivered from engineered cells  
 or gene therapy vectors.  
 (Updated on 25-MAR-2003 to correct PR field.)

SQ Sequence 3798 BP; 613 A; 1311 C; 1212 G; 662 T; 0 other;

Alignment Scores: 1.11e-233 Length: 3798  
 Pred. No.: 3505.00 Matches: 719  
 Score: 3505.00 Conservative: 122  
 Percent Similarity: 72.94% Mismatches: 260  
 Best Local Similarity: 62.36% Indels: 52  
 Query Match: 59.40% Gaps: 13  
 DB: 19

US-09-042-460-2 (1-1122) x AAV27876 (1-3798)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 Db 68 ATGCGGCGGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGGCAGCACTACCGCAG 127  
 QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
 Db 128 GTGTCGCGCTGCGCACGTTGCTGCGCGCTGCGGCGCCCGCGGCTGCGGCTGCGG 187  
 QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethIst 60  
 Db 188 CGCGGGACCCCGCGGCTTTCGCGCGCTGCTGCGCAGTGCCTGGTGGCGCTGG 247  
 QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 Db 248 GAGCACGCGCGCGCGCGCGCGCGCTCTCTCCGCGAGTGTCTCTGCTGAAGAGCTG 307  
 QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 308 GTGCGCGAGTGTGCGAGAGCTGTGCGAGCGCGCGCGCGAGAACGTGTGCGCTCGC 367  
 QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 Db 368 TTCGCGCTGTGAGCGCGCGCGCGCGCGCGCTTACACCGCTTACACCGCGCTG 427  
 QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyValTrpMetLeuLeu 140  
 Db 428 AGCTACCTGCCCAACACGCTGACCGAGCTGCGGGGAGCGGGGCGGGGCGGCTGCTG 487  
 QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 488 CTGCGCGCGTGGCGGCGGCGGCTGTGTTCACCTGTGCGCAGCTGCGGCTCTTTG 547  
 QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 Db 548 CTGTGGCTCCAGCTCGCTTACAGGTGTGCGGCGCGCGCTGTACCGCTTCGCGCT 607  
 QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 Db 608 GCCACTCAGCGCGCGCGCGCGCGCTAGT---GGACCCCGNAGCGCTCTGGGA--- 661  
 QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 Db 662 -----TGCGAACGGCGCTGGACCATAGCTCAGGAGCGCGG 700  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 701 GTCCCGCTGGCGCTGCGCGCGCGCGGTGCGAGAGCGCGGGGCGAGTCCCGCAAGT 760  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 761 CTGCGTGTCCCAAGAGCGCGCGCGCTGCGCGCTGCGCGCGGAGCGCGCGCTT 820  
 QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
 Db 821 GGGCAGGGCTCTCGTCTTCTGT 880  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 881 GTGGTGTCACTGCGCAGA-----CCCGCGGAGAGAGCGCACCTCTTTGGAG 925  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311

[illegible]

1994 AGAAGCTTCCGACAGAAAAGAGGGCGCAGCGTCCTCACTCGAGGGTGAAAGGCACGTGTT 2056

653 SerMetLeuAsnTyrlGluArgThrLysHisProHisLeuMetGlySerValLeuGly 672  
: : : : :  
2054 ACGGTGCTCAACTACGAGGCGCGCGCGGCCCTCTCTGGGCGCCTCTGTGTGGC 2113

673 MetAsnAspIleTyArgGThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
: : : : :  
2114 CTGGACGATATCCACAGGGGCTGGCGCACTTCGTGCTGCGTGTGGGGCCCACGAACCG 2173

693 ThrProArgMetTyrlPheValLysAlaAspValThrGlyValaTyAspAlaIleProGln 712  
: : : : :  
2174 CCGCTGAGCTGACTTTGTCAAAGTGGATGTAGCGGGCGGTACGACCATTCCCCCAG 2233

713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrcyAlle 732  
: : : : :  
2234 GACAGGCTCACGGAGGTCATCGCCACGATCATCAAA--CCCAAGAACGCTACTGTGGTG 2290

733 ArgGlnTyralaValArgAspSerGinglInValHisLysSerPheArgArg 752  
: : : : :  
2291 CGTGGTATGCGTGTCTCAGAAGGCCGCCATGGCAGCTCCGAAGGCTTCAAGAGC 2350

753 GlnValThrLysLeuAspLeuGlnProTyrlMetGlyGlnPheLeuLysHisLeuGln 772  
: : : : :  
2351 CACGCTCTACTCTGCAGACAGCCTCAGCGGTATCATCGCACGTTCGTGGCTCACCTGCAG 2410

773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluglnSerIleSerMetAsn 792  
: : : : :  
2411 GAAGCC-----AGCCGCTGAGGGATGCCGTCTCATCGACAGAGCTCTCCCTGNAT 2464

793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValVallys 812  
: : : : :  
2465 GAGGCGACGAGTGGCTCTTCGACGCTTCCTACGCTTCATCGCTTCATGTGCACACGCGCTGGC 2524

813 IleGlyAspArgCysTyrlThrGlnCysGinglYlleProGlnglySerSerLeuSerThr 832  
: : : : :  
2525 ATCAGGGCAGAGTCTTAGTCCAGTGGCAGGGATCCCGAGGGCTCATCTCTCTCCACG 2584

833 LeuLeuCysSerLeuCysPheGlyVasPmetGluAsnLysLeuPheAlaGluValGlnArg 852  
: : : : :  
2585 CTGCTCTCGACCTGTGTACGGCGACATGAGAAACAAGCTGTTCGGGGATTTCGGCGG 2644

853 AspglyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHileuAspGln 872  
: : : : :  
2645 GACGGCTGCTCTCTGCGTGTGGTGGATTTCTTGTGGTGCACCTCACCTCACCCAC 2704

873 AlaLysThrPheLeuSerThrLeuValHisglYalProGluTyrgLyCysMetIleAsn 892  
: : : : :  
2705 GGAAAACTTCTCTCAGAACCTTGCTCGAGGTGTCCCTGAGTATGGCTGGGTGGTGAAC 2764

893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlycylAlaalaPro 912  
: : : : :  
2765 TTGGGAGAGACAGTGGTGAACCTTCCTGTATAAGAGAGAGCGCCCTGGGTGGACGCGTTTT 2824

913 TyrlGlnLeuProAlaHisCysLeuPheProTrpCysglyLysLeuLeuAspThrGlnThr 932  
: : : : :  
2825 GTTCAGATCGCGGCCACCGGCTATTCCCTGTGTGGGCTGTGCTGTGATACCCGGACC 2884

933 LeuGluValPheCysAspTyrlSerGlyTyrlAlacInThrSeriLysThrSerLeuThr 952  
: : : : :  
2885 CTGGAGGTGCAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACC 2944

953 PheGlnSerValPheLysAlaglyLysThrMetArgAsnLysLeuLeuSerValLeuarg 972  
: : : : :  
2945 TTCACCGCGGCTTCACAGGCTGGAGGAAATGCGTCGCAAACTCTTGGGTCTTGGCG 3004

973 LeuLysCysHisgilyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCyale 992  
: : : : :  
3005 CTGAAGTGTACAGGCTTTTCTGGATTTCAGAGGTGAACAGGCTCCAGAGGTTGCACC 3064

993 AsnileTyrlLysIlePheLeuLeuGlnAlaTyrlArgPheHisAlaCysvalIleGlnLeu 1012  
: : : : :  
3065 AACATCTACAGATCCTCTGCTCAGCGGTACAGTTTCAGCATGTGTGTCGAGCTC 3124



533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTCGACCTGGCTGATGAGTGTGTACGTGCTGGTCAGCTGCTCAGGTCTTCTTTTAT 1741  
553 11eThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTCACGGAGACCAGTTTCAAAAGACACGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgLeuLeuSer 592  
1802 AAGTTGCAAGACATTGGAATCAGACAGCACTTGAAGAGGGGTGCAGCTCGCGAGAGCTGTCG 1861  
593 GlnGluGluValArgHisLysAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACAGAGGTTCAGCGAGCATCGGGAAGCCAGGCCCGCCTGCTGACGTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCCTGACGGGCTCGCGCGCATGTGTGAACATGACATGCTCGTGGGAGCC 1981  
633 ArgAlaLeuGlyArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTCCGACAGAAAGAGGGCGGAGCGCTCTCACTCGAGGGGTGAGGCACGTGTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCTCAATCAGACGGCGCGCGCCCGCCTCTGGGGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGGCTTGGCGCACCTTCGTGCTGCGGGCCCGACGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGTACTTTGTCAAGGTGATGTAGCGGGCGGTAGCACCATCCCCCAG 2221  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAAA--CCCCAGAACACGTACTGCTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGGTATCGCTGGTCTCCAGAAGCGCGCCATGGGCACGTCCGCAAGGCCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCCTACTTGCAGACCTCCAGCGGTATACGCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----ACCCGCGTGGGGATCGCGTGCATCGACAGAGTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCCAGCTGSCCTCTTCGACGCTTCTTACGGTTCATGTGCCACCACCGCGGTGCGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTCTACGTCCAGTGCAGGGGATCCCGAGGGTCTCATCTCTCTCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTCAGCCCTGCTACGCGCAGCATGGAGAAACAAGCTGTTTCGGGGATTTCGGCG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTGCGCTTGTGGTGGAGATTCTTGTGTGACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GGGAAAATCTTCTTCAGCACTCTGGTCCGAGGTGTCCTCGAGTATGAGTGTGCTGTGAGAC 2752

```
893 LeuGlnLysThrValAlaAspPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
|||||
2753 TTGCGAAGACAGTGGTGAACCTCCCTGTAGAACAGAGCCCTGGGTGGCAGCGCTTTT 2812
|||||
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
|||||
2813 GTTCAGATGCCGGCCACCGGCTATTCCCTGTGTGGCCCTGTCTGTGGATACCGGACC 2872
|||||
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerLeuLysThrSerLeuThr 952
|||||
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATATCCCGGACCTCCATCAGAGCAGCTCACC 2932
|||||
953 PheGlnSerValPheLeuAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
|||||
2933 TTCACCGCGGCTTCAGGCTGGAGGACATGTGTGCAACCTCTTGGGTCTTGGGG 2992
|||||
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
|||||
2993 CTGAAGTGTGCAGAGCTGTTCTGATTTGCAGGTGAACAGCTCCAGACGGTGTGCACC 3052
|||||
993 AsnIleTyrLysLeuPheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
|||||
3053 AACATCTCAAGATCTCTCTGTGAGCGGTACAGTTTCAGCATGTGTGTGTCAGCTC 3112
|||||
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
|||||
3113 CCATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTTGGCGTCACTCTGACACG 3172
|||||
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
|||||
3173 GCTCTCCTCTGTCTACTCCATCTCAAGCCAGAACGAGGATGTCTGTGGGGCCAG 3232
|||||
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
|||||
3233 GGGCGCGCGGCTCTGCTCCGAGCGCGGAGTGTGTGTCACCAAGCATTCCTG 3292
|||||
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
|||||
3293 CTCAGCTGACTCGACACCGGTTCACCTACGTGCCACTCTCTGGGTCACTCAGACAGCC 3352
|||||
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
|||||
3353 CAGACGAGCTGAGTGGAGTCTCCGGGGAGCAGCTGACTGCCCTGGAGCGCGAGCC 3412
|||||
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
|||||
3413 AACCGCGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3451
|||||
```

SULT 11  
Z30154

AAZ30154 standard; cDNA; 4015 BP.

AAZ30154;

26-JAN-2000 (first entry)

cDNA encoding a human telomerase reverse transcriptase (TERT).

Human; telomerase reverse transcriptase; TERT; T lymphocyte activation;  
dendritic cell; telomerase activity; cancer cell; proliferating cell;  
immunological destruction; telomerase; cancer; proliferation disease;  
ss.

Homo sapiens.

Key Location/Qualifiers  
CDS 56..3454

/\*tag= a

/product= "telomerase reverse transcriptase"

W09950392-A1.

07-OCT-1999.

```
XX 30-MAR-1999; 99WO-US06898.  
PF |||||  
XX 31-MAR-1998; 98US-0112006.  
PR |||||  
XX (GERO-) GERON CORP.  
PA |||||  
XX Gaeta FCA;  
PI |||||  
XX WPI; 1999-610845/52.  
DR P-PSDB; AAY43621.  
XX |||||  
PT Eliciting an in vivo immune response for prevention and treatment of  
PT cancers -  
XX |||||  
PS Disclosure; Fig 2; 26pp; English.  
XX |||||  
CC The present sequence encodes a human telomerase reverse transcriptase  
CC (TERT) polypeptide. The protein is used in the method of the invention.  
CC The specification describes a method for activating a T lymphocyte,  
CC comprising contacting the T lymphocyte with a dendritic cell that  
CC expresses a TERT peptide in the context of a MHC class I or MHC class  
CC II molecule. The protein causes induction of an in vivo immunological  
CC response to telomerase activity. Cancer cells are characterized by  
CC expression of endogenous TERT gene and the presence of detectable  
CC telomerase activity. Therefore, by eliciting a specific immune response  
CC to TERT or to TERT-expressing cells, it is possible to selectively target  
CC proliferating cells for immunological destruction. The method is used  
CC for eliciting an in vivo immune response to telomerase by activating  
CC a T lymphocyte, and is useful for prevention and treatment of cancers and  
CC other proliferation diseases/conditions.
```

XX Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

Alignment Scores:  
Pred. No.: 1.2e-233 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13

US-09-042-460-2 (1-1122) x AAZ30154 (1-4015)

```
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
|||  
DB 56 ATGCGCGCGCTCCCGCTGCGAGCGCGTCCCTGCTGCGCAGCCACTACCCGCGAG 115  
|||  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
|||  
DB 116 GTGCTGCGCTGGCCACGTTGCTGCGCGCTGGGGCCCGAGGGCTGGCGGTGTGCGAG 175  
|||  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
|||  
DB 176 CGCGGGGACCCCGCGGCTTTCGCGCGCTGTGGCCCGAGTGCCTGTGTGCGCTGTG 235  
|||  
QY 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
|||  
DB 236 GAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295  
|||  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
|||  
DB 296 GTGGCGCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
|||  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
|||  
DB 356 TTCGGCGTGTGTGAGCGGGCGCGCGGGGGCGCGCGGGCGCGCGCGCGCGCGCGCGCG 415  
|||  
QY 121 SerTyrLeuProAsnThrValIleLeuThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
|||  
DB 416 AGCTACCTGCCCAACACGCTGACGACGCACTGGCGGGGAGCGGGCGGTGGCGGTGCTG 475  
|||  
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
|||
```

476 CTGCCCGGCTGGGCGACAGCTGCTGCTTCACTGCTGGACGCTCGCGCTCTTTGG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGCTGGCTACACAGGTGGCGGCGCGCTGTACACAGCTCGCGCT 595  
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCACCTCAGCGCGCGCGCGCGCACACGCTAGT---GGACCCCGAAGCGCTCGGGA--- 649  
201 AsnPheThrAsnLeuPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGCAGACGGCGCTGGAAACCATAGCTCGCGAGGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240  
689 GTCCCGCTGGCGCTCCAGCGCGGCTGCGAGGCGCGCGGCGAGTGCCACCGCAAGT 748  
241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
749 CTGCCGTTGCCAAGCGCGCGAGCGTGGCGCTGCCCTGAGCGCGAGCGCGCGGT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGGTCTCTGGCGCCACCGCGGCGAGCGGTGGACCGAGTGACCGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSer 292  
869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGAGCCACTTTTGGAG 913  
293 GlyLysValSerAspLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGGCGCTCTTGGCAGCGCGCACTCCACCCATCCGTGGCGCGCCAGCACCGCGGGC 973  
312 SerThrSerLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCATCCACATCCGCGCGCACCGTCCCTGGGACAGCGCTTGTCCCGGTGTAGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTCTCTACTCTCAGCGCAC---AAGGAGCAGCTGCGCGCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCAGCGCGCGAGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTCAGGCGCTGATGCCAGGACTCCCGCAGGTTGCCCGCGCTGCCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGCGCAATCGCGCGCTGTTCTGAGCTGCTGGGAACCAACGCGCAGTGCCTTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAGACGCACTGCCCGTGCAGCTGG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCCGTGTCTGTGCCGCGGAGAACCCCGAGGCTCTGTGGCGCGCCCGGAGGAGGAC 1381  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGGCTGGTGTGCTGCTCCGCGCAGCACAGACCCCTGCGAGGTGATC 1441  
453 GlyPheLeuAlaCysLeuCysLysValSerAlaSerLeuLeuTrpGlyThrArgHis 472  
1442 GCCTTCGTGGCGCGCTGCTGCGCGGCTGTGCGCGCGCTCTCGGGCTCCAGCGAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysPheIleSerLeuGlyLysTyrGlyLys 492

1502 AACGAACGCGCTTCTCAGAACACCAAGAGTTTCATCTCCCTGGGAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGAGGAGCTGACGTGGAAGATGAGCGTGGGACTGCGCTTGGCTGGCGAG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCACAGGGTGGCTGTCTCCGCGCGCAGACCGCTCTCGGTGAGGAGATCTGGGC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGCTGATGAGTGTACGTCTGCTGAGCTGCTAGGTCTTTCTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGAGACCGTTCNAAGAACAGAGCTCTTTTCTACCGAGAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGATCAGACGACCTTGAGAGGGTGCAGCTGGCGAGCTGTGC 1861  
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGGTTCAGGACGACATCGGAGAACCGAGCGCGCTCTGTGACGTCCAGACTCCG 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTTCATCCCAAGCTGACGGCTGCGCGCGATTGTGAACATGAGCTACGTCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGCGAGAGAAAGAGCGCGAGCTTCACCTCGAGGGGTGAAGCAGCTGTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCTGGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCAGAGGCTGCGCGCACCTTCTGCTGCTGCGGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTCTATCGCGCAGCATCATCAAA---CCCGAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCCGCTGGTCCAGAGCGCGCGCATGGGCGACGTCCGCGAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTTACCTTGACAGACCTCCAGCGGTACATCGCAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGGATCGCGTCTCATCGACGAGAGCTCTCCCTCGAAT 2452  
793 GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGAGTGGCTCTTCGACGCTTCTTACGCTTCATGTGCCACCGCGCGTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTACGTCCAGTGCAGGGATCCGCGAGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACGCTGTCTACGCGGACATGAGAACAGCTGTTTCCGGGGATTCGGCGG 2632



853	AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln	872	FT	CDS	56...3454
2633	GACGGGCTCTCTCGTTTGGTGGATGATCTTCTTTGGTGGACACCTCACCTCACCCAC	2692	XX		/*tag= a
873	AlaLeuThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn	892	XX		
2693	GGGAAACCTTCTCAGGACCTGTGTCGAGGTGTCCTGATGATGCTGGTGGTGAAC	2752	XX		
893	LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro	912	XX		
2753	TTGCGGAGACAGTGTGAACCTCTCTAGAGACGAGGCGCTGGTGGCGGCTTTT	2812	XX		
913	TyrGlnLeuProAlaHisCysLeuPheProTyrPysGlyLeuLeuAspThrGlnThr	932	XX		
2813	GTTACAGATCGCGGCCACCGGCTATTCCTCTGGTGGCGCTGCTGGATACCCGACC	2872	XX		
933	LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr	952	XX		
2873	CTGGAGTGCAGGCGACTACTCCAGCTATGCCGACCTCCATCAGAGCCAGTCTCACC	2932	XX		
953	PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg	972	XX		
2933	TTCAACCGGGCTTCAAGCTGGGAGCAATGCGTCCCAACTCTTTGGGGCTTTCGGG	2992	XX		
973	LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle	992	XX		
2993	CTGAAGTGTACAGCTGTCTTGTGATTTGAGTTGAGGTGACAGCTCCAGCGTGTGCACC	3052	XX		
993	AsnIleTyrIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu	1012	XX		
3053	AACATCTAAGATCTCTCTGTCGAGCGGTACAGTTTCAGCATGTGTCTCAGCTC	3112	XX		
1013	ProPheAspGlnArgValArgLysAsnLeuThrPheLeuGlyIleIleSerSerGln	1032	XX		
3113	CAATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTTGGCGGTATCTTGACAGC	3172	XX		
1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer	1052	XX		
3173	GCCTCCCTCTGCTACTCCATCTCTGAAGCAAGAACGAGGAGTGTCTGGGGGCCAAG	3232	XX		
1053	GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu	1069	XX		
3233	GGCGCGCGCGGCTCTGCTCCCGAGCGCGTGCAGTGTGTGCCCAAGCATTCCTG	3292	XX		
1070	LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla	1089	XX		
3293	CTCAAGTGTACTCGACACCGGTGTACCTACGTGCGCACTCTTGGGGTCACTCAGGACGCC	3352	XX		
1090	GlnIysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla	1109	XX		
3353	CAGACGAGCTGAGTCGGAAGCTCCCGGGAGCAGCGCTGACTGCTGGAGCGCGACCC	3412	XX		
1110	AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp	1122	XX		
3413	AACCGGCACTGCGCTCAGACTTCAGACCATCTCTGGAC	3451	XX		
17-JAN-2000	(first entry)				
Human telomerase reverse transcriptase (hTERT) cDNA.					
Telomerase reverse transcriptase; human; hTERT; cell proliferation;					
cancer; ss.					
Homo sapiens.					
Key	Location/Qualifiers				

FT	CDS	56...3454
XX		/*tag= a
PN	WO9950386-A2.	
XX		
PD	07-OCT-1999.	
XX		
PF	31-MAR-1999; 99WO-US07097.	
XX		
PR	31-MAR-1998; 98US-0052864.	
XX		
PR	03-AUG-1998; 98US-0128354.	
XX		
PA	(GERO-) GERON CORP.	
XX		
PI	Morin GB;	
XX		
DR	WPI; 1999-610842/52.	
XX		
DR	P-PSDB; AAY32090.	
XX		
PT	New catalytic polypeptide and polynucleotide, useful for increasing	
XX		
PT	catalytic activity in a cell	
XX		
PS	Disclosure; Fig 2; 24pp; English.	
XX		
CC	This is the nucleotide sequence of cDNA encoding human telomerase	
CC	reverse transcriptase (hTERT, see AAY32090). Human telomerase is a	
CC	target for diagnosing and treating diseases relating to cell	
CC	proliferation and senescence, such as cancer, or for increasing	
CC	the proliferative capacity of a cell. A claimed method for	
CC	increasing the proliferative capacity of a vertebrate cell,	
CC	especially a human or other mammalian cell, involves introducing	
CC	into the cell a recombinant hTERT polynucleotide encoding an hTERT	
CC	variant in which residues 192-323, 200-323, 192-271, 200-271,	
CC	222-240, 415-450, 192-323 and 415-450, or 192-271 and 415-450 are	
CC	deleted. A claimed method for reducing telomerase activity in a	
CC	cell involves introducing a recombinant polynucleotide encoding	
CC	an hTERT variant having a deletion of amino acids 192-450, 560-565,	
CC	637-660, 638-660, 748-764 or 1055-1071. The polynucleotides are	
CC	obtained by mutagenesis of the hTERT coding sequence.	
XX		
SQ	Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;	
Alignment Scores:		
Pred. No.:	1,2e-233	Length: 4015
Score:	2505.00	Matches: 719
Percent Similarity:	72.94%	Conservative: 122
Best Local Similarity:	62.36%	Mismatches: 260
Query Match:	59.40%	Indels: 52
DB:	20	Gaps: 13
US-09-042-460-2 (1-1122) X AAZ20279 (1-4015)		
Qy	1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu	20
Db	55 ATGCGCGCGCTCCCGCTGCGAGCGCTGCTGCTGCTGCGAGCCACTACCCGAG	115
Qy	21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluArgArgLeuValGln	40
Db	116 GTGCTGCGCTGCGACGCTTGTGCGGCGCTGGGCGCCGAGGCTGGCGCTGGTGAG	175
Qy	41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp	60
Db	176 CGCGGGACCGCGCGCTTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	235
Qy	61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
Db	236 GACGACG	295
Qy	81 ValAlaArgValValGlnArgLeuCysGlnArgAsnGluArgAsnValLeuAlaPheGly	100
Db	296 GTGGCCGAGTGTCTCAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG	355
Qy	101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120

356 TTGCGCTGTGTGACAGGGGCGCGGGGCGCCCGGAGGCGCTTTCACCAACGCGGTGCGC 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrrpMetLeuLeu 140  
416 AGTACTACTGCCAACACGGGTGACCGACGACATGCGGGGAGCGGGGGGTGGGGGCTGCTG 475  
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGCGCGCGCTGGGCGCACGACGTCTGGTTACCTGCTGGCGACGCTCGCGGCTCTTTGTG 535  
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGCTCGCCTACCAAGGTGTGCGGCGCGCGCTGTACCAAGCTCGCGCGT 595  
181 ThrThrAspIleTrrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCGCGGCGCCCGCCACACGCTAGT--GGACCCCGAAGCGCTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
650 -----TCCGAACGGGCGCTCGAACATACGTCGCGAGGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
689 GTCCCGCTGGCGCTGCAGCCCGGGTGCAGAGAGCGCGGGGGCAGTGCACGCCGAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
749 CTGCGGTGGCCAGAGGCCAGCGCTGCGCTGCGCTCCTAGCGCGAGCGACGCCCGT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrrp--- 272  
809 GGGCAGGGTCTGGGCGCCACCGGCGCAGAGCGCTGCACGAGTGCACGTGGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTCCACTGCCAGA-----CCGCGCGAAGAACGCCCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGGGTCTCTGGCAGCGCCACTCCACCCACTCGTGGGGCGCCGACACACGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCGCATCCACATCGCGGCGCACCACTGCTGGACACGCTTGTCCCGGTGTACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgLysAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTCTACTCTCCTCAGCGCAC--AAGGAGCAGCTCGCGCCCTCCTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCAGCCTGACTGGCGCTCGAGGCTGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTCTCCAGGCGCTCGATGTCAGAGGACTCCCGCGAGGTTCGCCCGCTCCCGCAGCGC 1210  
391 TyrTrrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATATCGGCGCCCTGTCTTCTGGAGCTGCTGGGACACGACGCGCTGCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCCTCAAGACGCACTGCGCGCTGCGAGCTGCG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCGTGTCTGTGCCGGAGAACCCAGGGCTGTGTGGCGGCGCCCGAGGAGGAGC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrrpGlnValTyr 452

Db	1382	ACAGACCCCCGTCGCTGGTGGCAGCTGCTCCGCCACGACACAGACGCCCTCGCAGGTGTAC	1441
Qy	453	GlyPheLeuArgAlaCysLeuCybLysValValSerAlaSerLeuThrPGLyThrArgHis	472
Db	1442	GGCTTCGTGGGGGCTGCCTGGCGCGCTGGTGGCCCCCAGGCTCTGGGCGCTCCAGGCAC	1501
Qy	473	AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys	492
Db	1502	AACGACGGCGCTCTCCAGGACACCAAGAGTTTCATCTCCCTGGGGAGCATGCCAAG	1561
Qy	493	LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrrtpLeuArgSer	512
Db	1562	CTCTCGCTGCAGGAGCTGACGTGGAAGATGACCGTGGCGGACTGGCTTGGCTGCGCAGG	1621
Qy	513	SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla	532
Db	1622	AGCCAGAGGGGTGGCTGTGTTCGGCGCGCAGACACCGCTCTGCTGGAGAGATCCTGGCC	1681
Qy	533	ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuLeuArgSerPhePheTyr	552
Db	1682	AACTTCCCTGCTGGCTGATGAGTGTGTACGTGCTGGAGCTGCTCAGGTCTTTCCTTTAT	1741
Qy	553	IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer	572
Db	1742	GTCCGAGAGACCAAGTTTCAAAAGAAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC	1801
Qy	573	LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer	592
Db	1802	AAAGTTGCAAGCATTTGAAATCACAGACACTTGAAAGAGGTGCAGCTCGGGAGCTGTGC	1861
Qy	593	GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg	612
Db	1862	GAAGCAGAGGTCAGCAGCATCGGGAAGCCAGCGCGCTCTGCTGACGTCCAGACTCCGC	1921
Qy	613	PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr	632
Db	1922	TTTCATCCCAAGCCTCAGCGGCTCGCGCGCATGTGTAACTGACTACGTGCTGGAGCC	1981
Qy	633	ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe	652
Db	1982	AGAACGTTCCGAGAAAGAGGGCGGAGCGCTCTCACTCGAGGGTGAAGCACTGTC	2041
Qy	653	SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly	672
Db	2042	ACGGTGTCTCAACTACGAGCGGGCGCGCGCGCGCTCTCGGGCGCTCTGTGCTGGCC	2101
Qy	673	MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln	692
Db	2102	CTGGACGATATCCACAGGGCTGGCGACCTTGTGTGTGCTGCTGGGGCCCGAGCCCG	2161
Qy	693	ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln	712
Db	2162	CGCCTGAGCTGTTACTTGTCAAGGTGGATGTGACGGCGCGGTACGACCATCCCCCAG	2221
Qy	713	GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle	732
Db	2222	GACAGGCTTACGAGGGTCAATCGCCACCATCATCAA---CCCCAGAACGCTACTGGGT	2278
Qy	733	ArgGlnTyrAlaValValArgAspSerGlnGlnValHisLysSerPheArgArg	752
Db	2279	CGTCGGTATGCGTGGTCCAGAGCGCGCCCATGGGACGTCCGCAAGGCTTCAAGAGC	2338
Qy	753	GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln	772
Db	2339	CACGCTCTACTTTCAGACACTTCCAGCGCTACATGCGACAGTTCGTGGCTACCTGCAG	2398
Qy	773	AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn	792
Db	2399	GAGACC-----AGCCGCTGAGGGATGCGTCTCATCGACAGAGTCTCTCCCTGAAT	2452
Qy	793	GluSerSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys	812
Db	2453	GAGGCGCAGTGGCGCTCTTCAGCTGTCTCTACGTTTCATGCTGCGACACGCGCTGCG	2512

813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 932  
|||  
2513 ATCAGGGCAAGTCTCAGTCCAGTCCAGGGATCCCGAGGCTCCATCTCTCCACG 2572  
|||  
833 LeuLeuCysSerLeuLeuPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 952  
|||  
2573 CTGCTCTGAGCTGTCTACGGGACATGAGAACAGCTGTTTGGGGGATTCGGCG 2632  
|||  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 972  
|||  
2633 GACGGGCTCTCTCGTGTGGTGGATGATTTCTTGTGGTGACACCTCACCCAC 2692  
|||  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 992  
|||  
2693 GCGAAACCTTCTCAGGACCTGTGCGAGGTGTCCCTGAGTATGCTGCTGGTGTAC 2752  
|||  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
|||  
2753 TTGCGGAGACAGTGTGTAATCTTCCCTGTAGAACGAGGCTTGGGTGGCAGGGCTTT 2812  
|||  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuAspThrGlnThr 932  
|||  
2813 GTTCAGATCCCGGCCACAGCCCTATTCCTCGTGGGCTCTGCTGATACCCGAC 2872  
|||  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
|||  
2873 CTGAGGTGCAGAGGACTACTCCAGCTATGCCGACCTCCATCAGAGCCAGTCTCAC 2932  
|||  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
|||  
2933 TTCAACCGGGCTTCAAGGCTGGGAGAACATGCTGCAACATCTTTGGGGTCTTCGG 2992  
|||  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
|||  
2993 CTGAAGTGTACAGCTGTCTTCTGATTTGAGGTGACACCTCCAGACGGTGTGCACC 3052  
|||  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
|||  
3053 AACATCTACAAGATCTCTCTGTGAGGCGGTACAGGTTTCAGCATGTGTGTCAGCTC 3112  
|||  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleLeuSerSerGln 1032  
|||  
3113 CATTTTCATCAGCAAGTTTGGAGAACCCACATTTTCTTGGCGGTCACTCTGCACAG 3172  
|||  
1033 AlaSerCysTyrAlaLysLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
|||  
3173 GCCTCCCTCTGTCTACTCTCTGAAAGCAAGAACGAGGATGTCTGCTGGGGCCCAAG 3232  
|||  
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
|||  
3233 GGGCGCGCGGCGCTCTGCTCTCCGAGGCGGTGAGTGTGCTGTGCGCACAGCATTCCTG 3292  
|||  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
|||  
3293 CTCAAGCTGACTGCACACCGTGTCACTACGTGCCACTCTCTGGGTGACTCAGACAGCC 3352  
|||  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
|||  
3353 CAGACGAGCTGAGTGGAGCTTCCCGGGGAGCAGCGTACTGCTGGAGCGCGCAGCC 3412  
|||  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
|||  
3413 AACCGGCACTGCGCTCAGACTTCAGACCATCTCTGGAC 3451  
|||  
SULT 13  
Z00724  
AAZ00724 standard; DNA; 4015 BP.  
AAZ00724;  
06-OCT-1999 (first entry)

Human telomerase catalytic domain DNA.  
Telomerase; catalytic domain; human; quantitation; tumour cell; melanoma;  
body fluid; metastases; T-cell lymphoblastoma; chronic myeloid leukemia;  
acute lymphatic leukemia; melanoma; pulmonary carcinoma; colon cancer;  
breast cancer; ss.  
Homo sapiens.  
DE19804372-A1.  
05-AUG-1999.  
04-FEB-1998; 98DE-1004372.  
04-FEB-1998; 98DE-1004372.  
(DAHM/) DAHM M W.  
Dahm MW;  
WPI; 1999-431408/37.  
Quantifying tumor cells by amplifying mRNA encoding the catalytic  
subunit of telomerase  
Example; Fig 1A-B; 26pp; German.  
This invention describes a novel method for the quantitation of tumour  
cells in a body fluid which comprises (1) enrichment or isolation of  
tumour cells in the sample, (2) amplification of mRNA from these cells  
that encodes the catalytic subunit of telomerase and (3) quantifying  
the amount of amplified mRNA. The method is applied to tumour cells  
derived from (micro)metastases, e.g. associated with a wide range of  
tumours such as T-cell lymphoblastoma, chronic myeloid or acute  
lymphatic leukemia, melanoma, pulmonary carcinoma, cancer of colon or  
breast etc. This sequence encodes a human telomerase protein catalytic  
domain.  
Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.2e-233 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13  
US-09-042-460-2 (1-1122) x AAZ00724 (1-4015)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGAGCCACTACCGCGAG 115  
QY 21 ValTyrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTCGCTGGCCACGCTTCTGCGCGCGCTGCGGCGCCAGCGGCTGCGCTGCGAG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CGCGGGAGCCCGCGGCTTCCGCGCGCTGCGGCGCGCTGCGCTGCGCTGCGCTGCG 235  
QY 61 GlySerGlnProProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GACGACG 295  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGCGCGCGAGTGTGACAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 355  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
|||

356 TTGCGCGCTGCTGGACGGGGCCCGCGGGGGCCCCCGGAGGCGCTTCCACCACGAGCGTGGC 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
416 AGTACTCTCCCAACACCGGTGACCGACACTGCGGGGGAGCGGGCGTGGGGCGTCTG 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGCGCGCGTGGCGACGACGCTGTGTGTTCACTGCTGGACGCTGGCGCTTTGTG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGTCTCCAGTCTCCAGTCTGCGGGCGCGCGCTGTACCACTGCGGCT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCACCTCAGCGCGCGCGCGCGCACGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
650 -----TGCGAAGCGGCGCTGGAAACCATAGCGTCAGGAGCGCGGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCTCTGGGCTGCGACCGCGGTGCGAGGCGCGGGCGGAGTGCCAGCGCGAAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
749 CTGCGCTTGCACAGAGCCAGCGGTGCGGCTGCCCTGAGCGGAGCGGACCGCGTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGCAGGGGTCTTGGGCCCCACCGCGGCGAGCGCGTGACCGAGTGAACGTTCTTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
869 GTGGTGTCACTTGCACGA-----CCGCGCGAAGACCACTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGCGCTCTCTGGCAGCGGCGCACTCCACCCATCCGTGGCGCGCACGACACGCGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCACATCGCGGCCACACGCTCTCTGGACACGCTTGTCCCCCGGTGACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGAC---AAGGAGCAGCTGCGGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCAGCTGACTGCGCTCGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCCCTGGATGCGAGGACTCCCCGACGTTGCCCGCGCTGCCCGACGGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTGTCTTCTGGAGTGTCTTGGAAACCAACGCGCAGTGCCTTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAGACGCACTGCCGCTGCGAGTGGC-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTGCCCGGAGAACCCAGGCGCTCTGTGGCGGCCCCCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGTGTGTGAGCTGTCTCGCCGACGACGACGACCGCCCTGCGAGGTGTAC 1441

QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTTCGTGGCGGCTCTGCTGGCGCGCTGGTGGCCCCCAGGCGCTCTGGGGCTCCAGGCAC 1501  
QY 473 AsnGluArgArgPhePheLeuLysAsnLeuLysPheLeuLysSerLeuGlyTyrGlyLys 492  
Db 1502 AACGAACGCGCTTCTCTCAGGAACACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTGCGAGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGluLeuAla 532  
Db 1622 AGCCACGCGGTTGCTGTGTTCCGCGCGCAGAGCACCGTCTGCTGAGGAGATCCTGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1682 AAGTTCTCTGCACTGGCTGATGAGTGTGTAGCTGTGCTGAGCTCTTCTCTTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTACGAGAGCACGTTTCAAAAGACAGGCTCTTTTCTACCGGAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AAGTTGCAAGCATTTGGAATCAGACAGCACTTGAAGAGGCTGAGCTCGGAGCTGTGC 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTGACGCGAGCATCGGAAGCCAGGCGCGCTGTGCTGCTCCAGCTCCAGACTCCG 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCATCCCAAGCTGACGGCTCGCGCGCATTTGTGAACATGACACTACGTCGCGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAAGTTCCCGACAGAAAAAGAGCGCGCGCTCTACCTCGAGGGTGAAGGCACCTGTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCTCACTACGAGCGGGCGCGCGCGCTCTGCGCGCTCTGTCGTGGCC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGACGATATCCACAGGCGCTGGCGCACCTTCTGCTGCTGCGGGGCCAGGACCG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGTACTTGTCAAGTGTGTGACGCGCGCGTACGACACCATCCCCCAG 2221  
QY 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGCTCAGGAGTCACTCCCGCAGCATCATCAA---CCCCAAGAACAGTACTCGGTG 2278  
QY 733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
Db 2279 CGTCGTATGCGGTGTCAGAAAGCGCGCCATCGGGCACGTCGCGAAGGCGCTTCAAGAGC 2338  
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2339 CACGTCTCTACTTTCAGAGACTCCAGCGGTACATGGACAGTTCGTGGCTCACTCGAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2399 GAGACC-----AGCCGCGTGGGATCGCGTCTCATCGAGCAGAGCTCTCCCTCGAAT 2452  
QY 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
Db 2453 GAGCCAGCAGTGGCTCTCTCGACGCTTCTCTAGCTTCTATGTGCCACACGCGCGTGC 2512

```
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
|||
2513 ATCAGGGGCAAGTCTTACATCCAGTCCAGGCGATCCCGAGGGGTCCATCTCTCCACG 2572
|||
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAenLysLeuPheAlaGluValGlnArg 852
|||
2573 CTGCTCTGAGCGCTGTGCTACGGCGACATGAGAACAGCTGTTTGGGGGATTCGGCGG 2632
|||
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872
|||
2633 GACGGGCTGCTCTCGGTTTGGTGGATGATTTCTTGGTGGACACCTCACCTCACCCAC 2692
|||
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
|||
2693 GCGAAGACCTTCTCAGGACCTGCTCCAGGTGTCCCTGAGTATGGTGGTGGTGAAC 2752
|||
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
|||
2753 TTGGGAAGACAGTGGTGAATCTCCCTGTAGAGACGAGGCCCTGGGTGGACCGCTTTT 2812
|||
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
|||
2813 GTTCAGATGCGGCGCCACGCGCTATTCCCTGTGGTGGGCTCTCTCTGGATACCGGACC 2872
|||
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
|||
2873 CTGGAGGTGCAGAGCGACTACTCCAGTATGCCGAGCCTCCATCAGAGCCAGCTCTCAC 2932
|||
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
|||
2933 TTCAACCGCGCTTCAAGGTGGGAGAACATGCTGTCGGAACCTTTTGGGGTCTTGGCG 2992
|||
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
|||
2993 CTGAAGGTGCAGCGCTGTCTTCTGGATTTCAGGTCAACAGCCTCCAGACGGTGTGCACC 3052
|||
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
|||
3053 AACATCTACAAGATCTCCCTGCTGCGAGCGGTACAGGTTTCAGCGATGTGTCTGCAGCT 3112
|||
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
|||
3113 CCATTTATCAGCAAGTTTGGAGAACCCACATTTTCTCGCGGTCTATCTCTGACAGC 3172
|||
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
|||
3173 GCTTCCTCTGCTACTCCATCTCTGAAAGCCAGAACGCGAGGATGTGCTGGGGCCCAAG 3232
|||
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
|||
3233 GCGCGCGCGCGCTCTGCTCTCGAGCGCGTGCAGTGGCTGTGCACCAAGCATTCCTG 3292
|||
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
|||
3293 CTCAGCTGTACTCGACACCGGTGTCCCTAGTGGCCATCTCTGGGGTCACTCAGGACAGCC 3352
|||
1090 GlnLysLeuLeuCysArgGlyLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
|||
3353 CAGACGAGCTGAGTGGAGAGCTCCCGGGAGCAGCGTGTACTCTGCTGGAGCGCGAGCC 3412
|||
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
|||
3413 AACCGGCACTGCGCTCAGACTTCAAGACCATCTGGAC 3451
|||
RESULT 14
AAH45901
D AAH45901 standard; DNA; 4015 BP.
C
X AAH45901;
X
X 06-SEP-2001 (first entry)
X
X Human hTERT gene.
X
```

```
XX Human; hTERT; telomerase; catalytic subunit; mRNA quantitation;
KW detection; beta-region; diagnosis; cancer; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH exon 1..274
FT /tag= a
FT /number= 1
FT 275..1628
FT /tag= b
FT /number= 2
FT 1629..1924
FT /tag= c
FT /number= 3
FT 1825..2005
FT /tag= d
FT /number= 4
FT 2006..2185
FT /tag= e
FT /number= 5
FT 2186..2341
FT /tag= f
FT /number= 6
FT complement (2309..2325)
FT /tag= g
FT /note= "primer SYC1076 (AAH45902) binding site"
FT complement (2311..2325)
FT /tag= h
FT /note= "primer SYC1118 (AAH45905) binding site"
FT 2342..2437
FT /tag= i
FT /number= 7
FT complement (2345..2374)
FT /tag= j
FT /bound_moiety= "hybridisation probe CS3"
FT /note= "AAH45908"
FT complement (2427..2456)
FT /tag= k
FT /bound_moiety= "hybridisation probe CS12"
FT 2438..2523
FT /tag= l
FT /number= 8
FT complement (2458..2487)
FT /tag= m
FT /bound_moiety= "hybridisation probe CS1"
FT 2489..2506
FT /tag= n
FT /note= "primer SYC1097 (AAH45904) binding site"
FT 2524..2637
FT /tag= o
FT /number= 9
FT 2615..2631
FT /tag= p
FT /note= "primer SYC1078 (AAH45903) binding site"
FT 2638..2709
FT /tag= q
FT /number= 10
FT 2710..2898
FT /tag= r
FT /number= 11
FT 2899..3025
FT /tag= s
FT /number= 12
FT 3026..3087
FT /tag= t
FT /number= 13
FT 3088..3212
FT /tag= u
FT /number= 14
FT
```

exon 3213..3350  
/\*tag= v  
/number= 15  
exon 3351..4015  
/\*tag= w  
/number= 16

BP1108789-A2.

20-JUN-2001.

15-DEC-2000; 2000EP-0127228.

16-DEC-1999; 99US-0465491.

(HOFF ) HOFFMANN LA ROCHE &amp; CO AG F.

Chang SP, Santini CD;

WPI; 2001-376930/40.

Quantitating expression of mRNA encoding hTERT, the catalytic subunit of telomerase, as an indicator of cancer, by amplifying RNA using primers complementary to hTERT gene sequence and quantitating amplified products

Claim 1; Page 5-7; 29pp; English.

The present sequence is that of the hTERT gene encoding the catalytic subunit of the human telomerase, comprising 16 exons, which is useful in a method for quantitating hTERT mRNA. The method is useful for detecting the presence of beta-region (a 182 nucleotide region consisting of exons 7 and 8) of the hTERT-mRNA in a human sample for diagnosis and prognosis of cancer. The method provides an accurate measure of telomerase activity by selectively measuring mRNA that encodes an active hTERT protein.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

## Alignment Scores:

ad. No.:	1.2e-233	Length:	4015
ore:	3505.00	Matches:	719
rcnt Similarity:	72.94%	Conservative:	122
st Local Similarity:	62.36%	Mismatches:	260
ary Match:	59.40%	Indels:	52
	22	Gaps:	13

-09-042-460-2 (1-1122) x AAH45901 (1-4015)

```

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
|||
56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAG 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
|||
116 GTGCTCCGCTGGCCACAGTTCGTGCGGCGCTGCGGCGCCCGAGGCGCTGGCTGCAG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
|||
176 CGCGGGGACCGCGCGCTTCGCGCGCTGTGTGGCCAGTGCCTGGTGTGGTCCCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
|||
236 GAGCAGCGCGCGCGCGCGCTTCCTCCGCGAGGTGCTCTGCTGAAGGAGCTG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
|||
296 GTGCGCCGAGTGTGCGAGAGGTGTGCGAGCGCGCGCGGAGAGAGTGTGCTTCGCG 355
101 PheGluLeuAsnGlnAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
|||
356 TTGCGCTGTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140

```

```

416 AGCTACTGCCCCAACACGCTGACGACGACCTGCGGGGAGCGGGGCGTGGGGCTGCTG 475
|||
141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
|||
476 CTGCGCGCGTGGCGGACGACGCTGGTTCCTGCTGGCAGCGCTGCGCGCTCTTTTG 535
|||
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
|||
536 CTGGTGGCTCCAGCTCCGCTACCAAGTGTGGGGCGCGCTGTACCACTGGCGCT 595
|||
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
|||
596 GCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
|||
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220
|||
650 -----TGGGAACGGGCTTGAACCATAGCTCAGGAGGCGCGG 688
|||
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerTrp 240
|||
689 GTCCCGCTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
|||
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
|||
749 CTGCGTGTGCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
|||
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272
|||
809 GGGCAGGGGTCTCTGGCGCCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
|||
273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerLys 292
|||
869 GTGTGTGTCACCTGCGCAG-----CCGCGCGAAGAACCCACTCTTTGGAG 913
|||
293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311
|||
914 GGTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
|||
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
|||
974 CCCCCATCCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
|||
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
|||
1034 GAGACCAAGCACTTCTCTACTCTCAGGGGAC---AAGAGCAGCTGCGCGCTCTTC 1090
|||
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeuIlePhe 370
|||
1091 CTACTCAGCTCTGTAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150
|||
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
|||
1151 CTGGTTCCAGCGCTCGATGCCAGGACCTCCCGCAGGTGGTGGCGCTGCGCGCGCG 1210
|||
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
|||
1211 TACTGGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270
|||
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
|||
1271 GGGGTGCTCTCAAGACGCACTGCCCGCTCGAGCTGCG-----GTCACCCCGACA 1321
|||
431 Leu-----Asn 432
|||
1322 GCCGTGTCTGTGCGCGGAGAACCCCGAGGCTCTGTGCGCGCGCGCGCGAGGAGGAC 1381
|||
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
|||
1382 ACAGACCCCGCTGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1441
|||
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
|||

```





telomerase reverse transcriptase; gene; ss.

Homo sapiens.

Key Location/Qualifiers

CDS 56..3454

/\*tag= a

/product= "Human telomerase reverse transcriptase"

WO200274948-A2.

26-SEP-2002.

21-MAR-2002; 2002WO-CA00378.

21-MAR-2001; 2001US-277811P.

(GERO-) GERON CORP.

Denning C, Clark AJ, Schiff JM;

WPI: 2002-759895/82.

P-PSDB; ABE29226.

Mammalian cells, useful for producing animal tissues with carbohydrate antigens that are compatible for transplantation into human patients - Disclosure; Page 33-34; 71pp; English.

The invention relates to animal tissues with carbohydrate antigens that are compatible for transplantation into human patients. The mammalian cell is inactivated homotypously for expression of alpha(1,3)galactosyltransferase (alpha1,3GT) gene and comprises a transgene for alpha(1,2)-fucosyltransferase (alpha1,2FT). It is useful for producing animal tissue with carbohydrate antigens that are compatible for transplantation into human patients. The present sequence is human telomerase reverse transcriptase (TERT) cDNA used in the invention.

Sequence 4015 BP; 663 A; 1363 C; 1275 G; 714 T; 0 other;

gment Scores:

id. No.:	1.2e-233	Length:	4015
re:	3505.00	Matches:	719
cent Similarity:	72.94%	Conservative:	122
it Local Similarity:	62.36%	Mismatches:	260
ry Match:	59.40%	Indels:	52
	24	Gaps:	13

09-042-460-2 (1-1122) x AAD46821 (1-4015)

```

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ARGCCGCGCTCCCGCTGCGAGCGGTGGCTCCCTGCTGGCGACCACTACCCGCGAG 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GFGCTGCGCTGGCCACGTTCTGTGGGGCTGGGGCCCGCAGGCTGGCGCTGGTGCAG 175
41 ProGlyAspProLeuValLeuValArgLeuValAlaGlnCysLeuValCysMetHisTrp 60
176 CCGCGGGACCCGCGCGCTTCGCGCGCTGTGGGCCAGTGCCTGGTGTGGTGGCTGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuValGlu 80
236 GACGACGCGCGCCCGCGCGCTTCCTCCGCGAGTGTCTGCTGAGAGAGCTG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGGCCCGAGTGTGACAGAGCTGTGGAGCGCGCGAGGAGAACGTCGTGGCTTCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGGAGCGGCGCGCGGGGGCCCCCGAGGCGCTTCACCAACCGGTGCC 415

```

```

QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
DB 416 AGCTACCTGCCCAACACGCTGACCGACGACCTGCGGGGAGCGGGGCGTGGGGCTGCTG 475
QY 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
DB 476 CTGGCGCGCTGGGCGAGCGAGCTGTGTTCACTGTGTCGACGCTGCGCGCTCTTTGNG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
DB 536 CTGGTGGCTCCAGCTGCGCTACAGGTGTGGGGCGCGCTGTGTACAGCTCGCGCT 595
QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
DB 596 GCCACTCAGCGCGCGCGCGCGCGCGCGCTAGT--GGACCCCGAGAGCGCTGTGGA--- 649
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220
DB 650 -----TGCAGAACGCGCGCTGGAACCATAGCTCAGGAGGCGCGG 688
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
DB 689 GTCCCTCTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
DB 749 CTGCGCTTGGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
DB 809 GGGCAGGGGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
DB 869 GTGTGTGTCACCTGCCAGA-----CCGCGCGAGAGACCCACCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
DB 914 GGTGCGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 974 CCCCATCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
DB 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGCGAC--AAGGAGCAGCTGCGCGCTCTTC 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
DB 1091 CTACTAGCTCTCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
DB 1151 CTGGTTCCAGGCGCTGGATGCGCAGGAGCTCCCGCAGGTTCGCCCGCTGCCCGCGCG 1210
QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
DB 1211 TACTGGCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
DB 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGGAGCTGCG-----GTACCCCGACGA 1321
QY 431 Leu-----Aen 432
DB 1322 GCGCGTGTCTGCGCGCGGAGAGCCCGCGCGCTCTGTGCGCGCGCGCGCGCGCGCGAG 1381
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
DB 1382 ACAGACCCCGCTGCTGTGTGAGCTGCTCCCGCAGCAGCAGCAGCAGCAGCAGCAGT 1441

```

453 GlyPheLeuArgAlaCysLeuCysLeuValValSerAlaSerLeuTyrGlyThrArgHis 472  
1442 GGCCTCGTGGCGCCTGCTGCGCGCTGGTGGCCCCCAGGCGCTTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLeuLeuTyrLysPheLeuLeuTyrGlyLys 492  
1502 AACGACCGCGCTTCTCAGGAACACCAAGAGTTCACTCCCTGGGGAGCATGCCAAG 1561  
493 LeuSerLeuGlnLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGTGCAGAGCTGAGTGGAGATGAGCGTCCGGGACTGGCGCTTGGCTGGCAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgLeuAla 532  
1622 AGCCAGGGGTGGCTGTGTTCGGGCGCGAGACCGCTCTGCGTGGAGAGATCCTGGCC 1681  
533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGCTGAGCTCTCAGGTCTTCTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572  
1742 GTCCAGGAGACCAAGTTTCAAAGAACAGGCTCTTTTCTACCGAAGAGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCAATGGATCAGACAGCACTTGAAGAGGCTGCAGCTGGGAGGTGTGCG 1861  
593 GlnGluValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgGluArg 612  
1862 GAAGCAGAGTTCAGGAGCATCGGAGAGCCAGGCGCGCTCTGCTGAGCTCCAGATCCCG 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCTGACGGCTGCGCGCGATTGTGAACATGACTACGTCTGTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCGCGAGAGAAAGAGGGCGGCGCGCTCTGCTGCGCGCTCTGTGTGGCG 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCAACTACGAGCGGCGCGCGCGCTCTGCTGCGCGCTCTGTGTGGCG 2101  
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCGCTGCGCACCTTCGTCTGCTGCGCGCGCGCGCGCGCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGTGTGATGTACGGGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 CACAGGCTCAGGAGGTTCATCGCGCAGCATCAAA---CCCCAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGATGCGGTGGTCCAGAGCGCGCCATGGGACGTCGCGAGGCCCTTCAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACTTGCAGACCTCCAGCGGTACATGCGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetHis 792  
2399 GAGACC-----AGCCCGCTGAGGAGTCCGTCTCATCGACAGAGCTCTCCCTGAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCCAGCAGTGGCTCTTCGAGCTCTTCCTACGCTTCATGTGCCACCGCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832

2513 ATCAGGCGCAAGTCTTACGCTCCAGTGCAGGGGATCCGCGAGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTCGAGCCTGTGTACGGCGCATGAGAACAGCTGTTTGGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCGCTTGGTGGATGATTTCTTGTGTGACACCTCACCCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTCTCTCAGGACCTGCTCGAGGTTCCTGAGTATGGCTCGCTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTCCGAGACAGATGGTGAATTCCTGTAGACAGGAGCCCTGCTGCTGGATACCCGACC 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCCGGCCCGCCAGCGCTATTCCTGTGGTGGCTGCTGGATACCCGAGC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGCAGAGCGACTACTCCAGCTATCCCGGACCTCCATCAGAGCGAGTCTACC 2932  
953 PheGlnSerValPheLeuAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGCGCTTCAAGGCTGGGAGGAACATGCTCGCAACTCTTTGGGCTCTTGGC 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCGCTGTTCTGGATTGAGGTGAACAGCCTCCAGAGCGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCTGAGCGGTACAGGTTTACGCGATGTGTGTGCAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTTCATCAGCAAGTTTGAAGAACCCCATTTTCTCGCGCGTCATCTCTGCACAGC 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCTCTGTACTTCCATCTCTGAAGCCAGAACCGAGGATGTGCTGGGGGCCAAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGCGCTCTGCTGCTGCGAGCGCTGCGAGTGGCTGTGCGCCACCAAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTCAGTTCGACCGGTACCTACGTGCGCATCTCTGGGGTCACTCAGGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGCACTGAGTGGAGGCTCCCGGGGACGAGCTGACTGCTGCGCTGGAGCGCGAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGCGCACTCCCTCAGACTTCAAGACCATCTCTGGAC 3451

Search completed: February 1, 2004, 15:20:47

Job time : 880 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

on: February 1, 2004, 15:05:59 ; Search time 144 Seconds

(without alignments)

3439.110 Million cell updates/sec

le: US-09-042-460-2

fect score: 5901

quence: 1 MTRAPRCFAVRLSRVRE.....TILKAADPALSTDPQTILD 1122

ring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

urched: 569978 seqs, 220691566 residues

al number of hits satisfying chosen parameters: 1139956

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

mmand line parameters:

>DEL=frame\_p2n.model -DEV=xlh

>/cgn2\_1/USPTO.spool/US0904260/runat 01022004 150552 17040/app.query.fasta\_1.1287

>Issued\_Patents\_NA\_QFMT=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

>OPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi

>LST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

>DE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

>SER=US0904260 @CGN 1 1.71 @runat 01022004 150552 17040 -NCPUS=6 -ICPU=3

>WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG

>V TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6

>AFEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

abase :

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Query Match	Score	Length	ID	Description
1	3505	59.4	3396	3	US-08-974-549A-638
2	3505	59.4	3396	3	US-08-974-549A-639
3	3505	59.4	3396	3	US-08-974-549A-640
4	3505	59.4	3396	3	US-08-974-549A-641
5	3505	59.4	3396	3	US-08-974-549A-642
6	3505	59.4	3451	3	US-08-974-549A-721
7	3505	59.4	4015	3	US-08-851-843A-224
8	3505	59.4	4015	3	US-08-974-549A-1
9	3505	59.4	4015	3	US-08-854-050-224
10	3505	59.4	4015	4	US-09-430-323-224
11	3505	59.4	4015	4	US-09-572-423B-3
12	3505	59.4	4015	4	US-09-128-354-1

Sequence 1, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 343, App  
Sequence 173, App  
Sequence 292, App  
Sequence 173, App  
Sequence 173, App  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 100, App  
Sequence 266, App  
Sequence 100, App  
Sequence 100, App  
Sequence 6, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 30, Appli  
Sequence 8, Appli  
Sequence 62, Appli  
Sequence 62, Appli  
Sequence 8, Appli  
Sequence 68, Appli  
Sequence 111, App  
Sequence 68, Appli  
GENERAL INFORMA  
Sequence 111, App  
Sequence 1, Appli  
Sequence 109, App

ALIGNMENTS

RESULT 1  
US-08-974-549A-638  
; Sequence 638, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/974,549A  
; APPLICATION NUMBER: 19-NOV-1997  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997

Prior Application Data:	
Application Number:	US 08/846,017
Filing Date:	25-APR-1997
Prior Application Data:	
Application Number:	US 08/851,843
Filing Date:	06-MAY-1997
Prior Application Data:	
Application Number:	US 08/854,050
Filing Date:	09-MAY-1997
Prior Application Data:	
Application Number:	US 08/911,312
Filing Date:	14-AUG-1997
Prior Application Data:	
Application Number:	US 08/912,951
Filing Date:	14-AUG-1997
Prior Application Data:	
Application Number:	US 08/915,503
Filing Date:	14-AUG-1997
Prior Application Data:	
Application Number:	WO PCT/US97/17618
Filing Date:	01-OCT-1997
Prior Application Data:	
Application Number:	WO PCT/US97/17885
Filing Date:	01-OCT-1997
Attorney/Agent Information:	
Name:	Apple, Randolph Ted
Registration Number:	36,429
Reference/Docket Number:	015389-002610US
Telecommunication Information:	
Telephone:	(415) 576-0200
Telex:	(415) 576-0300
Information for SEQ ID NO:	638:
Sequence Characteristics:	
Length:	3396 base pairs
Type:	nucleic acid
Strandedness:	single
Molecule Type:	DNA
Feature:	
Name/Key:	-
Location:	1..3396
Other Information:	/note= "hTET-encoding sequence employing alternative codon distributions for E. coli (all genes)"
Other Information:	E. coli (all genes)"
Alignment Scores:	
Seq. No.	0
Score:	3505.00
Percent Similarity:	72.94%
Best Local Similarity:	62.36%
Query Match:	59.40%
Gaps:	3
1-09-042-450-2 (1-1122) x US-08-974-549A-638 (1-3396)	
1 MetThrAlaProArgCysProAlaValArgSerLeuLeuArgSerGlyTrpArgGlu 20	
1 ATGCCCGCCGGCGCGCTGTGCAGCGCTGTGCAGCGCTGTGCAGCGCTATTCGCGAA 60	
21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProLuglIvargArgLeuValGln 40	
61 GTCGTCCGCTGGCGACCTTTTGCGCGCGCTGTGCAGCGCTGTGCAGCGCTGTGCAG 120	
41 ProGlyAspProLyseTyrrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60	
121 CGCGCGCATCCGGCGCGCTTTCGCGCGCTGTGCAGCGCTGTGCAGCGCTGTGCAG 180	
61 GlySerGlnProProAlaAspLeuSerPheHISGLNValSerSerLeuLysGluLeu 80	
181 GATCGCGCCCGCCCGCGCGCGCTTTCGCGCGCTGTGCAGCGCTGTGCAGCGCTGT 240	
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100	





814 GTTGTTCCTCCCGCTCGT-----CGGGCTGAAGACTACTCTCCCTGAA 858  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
859 GGTGTCTCTCCGGTACCGCTCACTCCACCGCTCCGGTGTCTGCAGCACCAACGCTGGT 918  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
919 CGCCCGTCCACCTCCGCTCCGGCTCGGGTCCGGTGGACACCCCGTCCCGCGGTTACGCT 978  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
979 GAAACCAACACTCTCTGTACTCTCCGCTGAC---AAAGAACAGCTGCGTCCGCTCTTC 1035  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370  
1036 CTGCTGTCTCTCCCTCGCTCCGCTCCGCTGACCGGTGCTCTGCTGTGTAACCACTCTTC 1095  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1096 CTGGTTCCCGTCCGTTGGATGCGGGTACCCCGCTGCTCTGCGCGCTCTGCGCAGCGT 1155  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1156 TACTGTCAGATGCTCCGCTGTCTCTGGAACCTGCTGGGTAAACCAACGCTCAGTCCCGTAC 1215  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1216 GGTGTCTCTGTAACCACTCTGCGCGCTGCTGCTGT-----GTTACCCCGCT 1266  
431 Leu-----Asn 432  
1267 GCTGTGTGTTCGCTCGTGAACACCGCAGGTTCCGTTGCTGCTCCGGAAGAAAGAC 1326  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1327 ACCGACCCGCTGCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386  
453 GlyPheLeuAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1387 GGTTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyTyrGlyLys 492  
1447 AACGAACGCTGCTTCTGCTGCTAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1506  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1507 CTGTCCTCGAGGACTGACTGGAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 TCCCGCGGTGTTGGTTCGCTCCGCTGCTGAACACCGCTCTGCGTGAAGAAATCCTGGCT 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1627 AANTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1687 GTTACCGAACCACCTTCCAGAAACACCGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AAATGTCAGTCCATCGGTATCGCTCAGCACCTGAAACCGTGTTCAGCTCGGTGAACCTGTC 1806  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAGCTGAAGTTCGTGACGACCGTGAAGCTCGTCCGGCTCTGCTGACCTCCCGCTCGGT 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632

1867 TTCATCCGAAACCGGACGCTCTCGCTCCGATCGTTAAATGACTAGCTGTGTGTGTGCT 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 552  
1927 CGTACCTTCCGCTCGTGAACACCGTGAACGCTGACCTCCCGTGTGTAAAGCTCTGTTC 1986  
653 SerMetLeuAenTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
1987 TCCGTTCTGAACACGACGCTCTCGTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 CTGACGACATCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2107 CGCGCGGACTGACTCTGTTAAGTTGAGTTACCGGTGCTTACGACACCATCCCGCAG 2166  
713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2167 GACCGTCTGACGGAAGTTATGCTTCCATCATCAAA---CCGACGAACACCTACTCGGT 2223  
733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 GCTGTTACGCTGTGTGTGTAAGAGCTGCTACGCTCAGCTTCTGTAAGCTTTCAATCC 2283  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CACGTTTCCACCTGACCGACCTGACCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2343  
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2344 GAAACC-----TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2397  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAAGCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2457  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATCCGCTGTAATCTAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2517  
833 LeuLeuCysSerLeuCysPhePheGlyAspMetGluAsnLysLeuPheAlaGluValArg 852  
2518 CTGCT 2577  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872  
2578 GACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2637  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2638 GCTAAACCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 CTGCTGTAACCGTGTGTTAATCTCCCGCTGTAAGACGAGCTCTGGGTGCTGCTGCTGCTGCT 2757  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAGATGCGGCTCAGGCTCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2817  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 CTGGAAGTTCTAGTCGACTACTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2877  
953 PheGlnSerValPheLeuAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGTGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 CTGAATGCCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2997



993 AsnIleTyrIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATCTACAAATCCCTGCTGCAGGCTTACCGTTCCACGCTTGCCTTCTGAGCTG 3057  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCGTTCACACGAGGTTTGAAAAACCCGACCTTCTTCTGCGTGTATCTCCGACACC 3117  
1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCCCTGCTACTCATCTGAAAGCTAAAAACGCTGTATGCTCCCTGGTGTCTAAA 3177  
1053 GlySer-----PheProGluAlaAlaHisIlePheLeuCysTyrGlnAlaPheLeu 1069  
3178 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 CTGAAACTGACCCGCTACCGTGTACCTAGCTTCCGCTGCTGCTGCTGCTGCTGCTGCT 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3298 CAGACCCAGCTGCTCCGTAACCTCCGCGTACACCTGACCCCTGCTGCTGCTGCTGCTGCT 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3396

RESULT 3  
5-08-974-549A-640  
Sequence 640, Application US/08974549A  
Patent No. 6166178  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 640:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3396 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: -  
LOCATION: 1..3396  
OTHER INFORMATION: /note= "HIRT-encoding sequence employing  
OTHER INFORMATION: alternative codon distributions for  
OTHER INFORMATION: yeast (all genes)"  
US-08-974-549A-640  
Alignment Scores:  
Pred. No.: 0 Length: 3396  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 13 Gaps: 13  
US-09-042-460-2 (1-1122) x US-08-974-549A-640 (1-3396)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
DB 1 ATGCCAAGAGCTCCAAAGATGTAGAGCTGTAGATCTTTGTTGAGATCTCATTTATAGAGAA 60  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
DB 61 GTTTTGCCATTGCTACTTTTGTAGAGATTGGGTCACAAAGGTGGAGATTGGTTCAA 120  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethIleTrp 60  
DB 121 AGAGTGATCCAGCTGCTTTAGAGCTTTGTTGCTCAATGTTTGGTTTGTTCATGG 180  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
DB 181 GATGCTAGACCAACCAGCTGCTCCATCTTTTAGCAAGTTTCTTTGTTGAAGAATTG 240  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
DB 241 GTTCTAGAGTTTGTGCAAGATTGTGTGAAGAGGTGCTAAAAATGTTTGGCTTTGGT 300  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
DB 301 TTTGCTTTGTTGATGTGTAGAGGTGTGTCACCAAGAGCTTTTACTACTTCTCTGTAGA 360  
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140

```
|||||
361 TCCTATTGCCAAATCTGTTACTGATGCTTTGAGAGGTTCTGGGCTTGGGTTGTTG 420
|||||
141 LeuSerArgValGlyAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
|||||
421 TTGAGAAGAGTTGGTGATGATGTTTGGTTCTATGTTGGCTAGATGCTTGTGTT 480
|||||
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnHisCysAla 180
|||||
481 TTGGTTGCTCCATCTTGTGCTTATCAAGTTTGTGCTCCACCATTTGATCAATGGGTGCT 540
|||||
181 ThrThrAspIleProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
|||||
541 GCTACTCAAGCTAGACACACCATGCTTCT--GGTCCAAAGAAGAAGATTGGGT-- 594
|||||
201 AspPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
595 -----TGTGAAGAGCTTGGAAATCATCTCTGTGAGAGAGCTGGT 633
|||||
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
634 GTTCCATTGGGTTTGCAGCTCCAGGTCGTAGAGAGAGGTGGTCTCTAGATCT 693
|||||
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
694 TTGCCATTGCCAAAAGACCAAGAAGAGGTGCTGCTCCAGAACCAAGAAGAACTCCAGTT 753
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
754 GGTCAAGGTTCTGGGCTCATCCAGTAGAAGCTAGAGGTCCTCATCTGATAGAGTTTGT 813
|||||
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
814 GTTGTGTTCTCCAGCTAGA-----CCAGCTGAAGAAGCTACTCTCTTGGA 858
|||||
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
859 GGTGCTGTTGCTGGTACTAGACATTCTCATCTGTTGTTGGTAGACAACATCATGCTGGT 918
|||||
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
919 CCACCATCTACTTCTAGACACCAAGACCATCGGATACCTCCATGCTCCACCATTTATGCT 978
|||||
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
979 GAAACTAAACATTTTGTATTCTCTGCTGAT---AAAGAACAAATTGAGACCATCTTTT 1035
|||||
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyValAlaArgArgLeuValGluIlePhe 370
1036 TTGTTGCTCTCTTGGAGACCATCTTTGACTGCTGCTAGNAGATTGTTGAACTATTTT 1095
|||||
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
1096 TTGGGTTTAGACCATGGATGCCAGGTACTCCAAAGAAGATTGCCAAGATTGCCAAAGA 1155
|||||
391 TyrTyrGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
1156 TATTGCCAAATGAGACCATTTTGGGAATTTGTTGGTAATCATGCTCATGTCATAT 1215
|||||
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
1216 GGTGTTTTTGTGAAGAACTCATTTGCTCCATTGAGAGCTGCT-----GTTACTCCAGCT 1266
|||||
431 Leu-----Asn 432
1267 GGTGGTGTGCTAGAGAAACCAAGGTTCTGTGCTGCTCCAGAGAAAGAGAT 1326
|||||
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452
1327 ACTGATCCAGAGAGATTGGTCAATTTGTTGAGACAACATTTCTTCCATGGCAAGTTAT 1386
|||||
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472
|||||
```

RESIT.T 4

```

Recent Similarity: 72.94% Conservative: 122
st Local Similarity: 62.36% Mismatches: 260
ery Match: 59.40% Indels: 52
: 3 Gaps: 13

-09-042-460-2 (1-1122) x US-08-974-549A-641 (1-3396)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
1 ATGCCAAGAGCTCCAGAGGTAGAGCTGTTAGATCTTGTGTGAGATCTCACTACAGAA 60
21 ValTrrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
61 GTTGTGCCATGTGCTCTTTCGTAGAGATGGGTCCACAGAGTTGGTTCAT 120
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrr 60
121 AGAGTGTACCCAGCTCTTCAGAGCTTGTGTCTCAATGTTGGTGTGTCCATGG 180
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerLeuLysGluLeu 80
181 GAGCTAGACACACAGCTGCTCCATCTTTCAGACAGATTTCTTGTTCAGAGATTTG 240
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
241 GTTGTAGAGTTTGTGAAGATTTGTGAAGAGGTGCTAGAGCTTTCGTTTCGGT 300
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
301 TTCGCTTGTTCAGCGTGTAGAGGTGTCCACAGAGCTTTCCTACTTCTGTAGA 360
121 SerTrrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrrMetLeuLeu 140
361 TCTTACTTGTCAACACTGTACTGACGCTTTCAGAGGTTCGTGTGGTGTGTGGT 420
141 LeuSerArgValGlyAspLeuLeuValTrrLeuLeuAlaHisCysAlaLeuTrrLeu 160
421 TTGAGAAGAGTGTGTGACGAGCTTGTGTCTTCTTGTGTGTGTGTGTGTGTGTGT 480
161 LeuValProSerCysAlaTrrGlnValCysGlySerProLeuTrrGlnIleCysAla 180
481 TTGCTGTCTCATCTTGTGCTTACCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
181 ThrThrAspIleTrrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
541 GCTACTCAAGCTAGACACACACACAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 594
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
595 -----TGTGAAGAGCTTGTGAACCACTCTGTGTAGAGAGCTGGT 633
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
634 GTTCCATTTGGTGTCCAGCTCCAGGTGTAGAGAGAGAGGTGTGTGTGTGTGTGTGTGT 693
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
694 TTGCCATTTGCCAAGACACCAAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 753
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrr--- 272
754 GGTCAGAGTTTCTGGGTCTCCAGGTGTAGAACTTAGAGGTTCATCTGTACAGAGTTTCTGT 813
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
814 GTTGTCTTCCAGCTAGA-----CCAGCTGAGAGAGCTACTTCTTTGGAA 858
293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311
859 GGTGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 918
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330

```



FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 642:

SEQUENCE CHARACTERISTICS:

LENGTH: 3396 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY:

LOCATION: 1..3396

OTHER INFORMATION: /note="generic" hTERT protein encoding

OTHER INFORMATION: sequence

-08-974-549A-642

Ignment Scores:

ed. No.: 0 Length: 3396

ore: 3505.00 Matches: 719

rcent Similarity: 72.94% Conservative: 122

st Local Similarity: 62.36% Mismatches: 260

ery Match: 59.40% Indels: 52

: 3 Gaps: 13

-09-042-460-2 (1-1122) x US-08-974-549A-642 (1-3396)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyArgGlu 20  
 1 ATGCCACGTGCCCCACCGTGTGCGGTCGTTCTTTGTTGCGTCTCTCACTACCGTCAA 60  
 21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgValGln 40  
 61 GTTTTCCCATGTGCCACCTTCGTTGCTGTTGGGTCCACAGGTGGCGTTGGTTCAA 120  
 41 ProGlyAspProLysIleTyArgThrLeuValAlaGlnCysLeuValCysMetHisTTP 60  
 121 CGTGTGTATCCAGCCGCTTCGTCGCTTGGTTGCCAATGTTTGTGTTGCCATGG 180  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 181 GATGCGCGTCCACCCACGCGCCGCGCCATCTTTCCGTCAGTTTCTTTGAAAGAATTG 240  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 241 GTTGGCCGTTTGTGCAAGTTTGTGACCTGTGTCGTCGTCGTCGTCGTCGTCGTCG 300  
 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 301 TTCGCTTGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 360  
 121 SerTyLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTTPMetLeuLeu 140  
 361 TCTTACTTGCACCAACACCGTTACCGAGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
 141 LeuSerArgValGlyAspLeuValTyArgValLeuLeuAlaHisCysAlaLeuTyLeu 160  
 421 TTGCGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 480  
 161 LeuValProProSerCysAlaTyArgGlnValCysGlySerProLeuTyArgGlnIleCysAla 180  
 481 TTGTTGCCCCCATCTTGTGCTACCAAGTTTGTGTGTCCACCATTCATGCAATTGGGTGCC 540

181 ThrThrAspIleTTPProSerValSerAlaSerTyArgProThrArgProValGlyArg 200  
 541 GCCACCCAGCCCGTCCACACACGCTCT---GGTCCAGTGTGTTGGGT---594  
 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 595 -----TGTAACGTGCGCTGGAAACCATCTCTGTCGTGAAGCGGT 633  
 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 634 GTTCATTTGGTTTGGCAGCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693  
 241 ValProSerAlaLysLysAlaArgCysTyProValProArgValGluGluGlyPro---259  
 694 TTGCCATTGCCAAACAGTCCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 753  
 260 -----HisArgGlnValLeuProThrProSerArgLysSerTTP---272  
 754 GGTCAAGGTTCTTGGGCCCCACCCAGGTCGTACCGTGGTCCATCTGATCGTGTTCGT 813  
 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 814 GTTGTTCCTCCAGCCCGT-----CCAGCCGAGAGACCCACTCTTTGGAA 858  
 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 859 GGTGCTTGTCTGTGTACCGGTCACCTCTCACCCATCTGTGTCGTCAACACACCGCGGT 918  
 312 SerThrSerLeuLeuSerProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 919 CCACCATCTACTCTGTCCTCACCGTCCTGATGATGATGATGATGATGATGATGATGAT 978  
 331 GluThrArgHisPheLeuTySerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 979 GAAACCAACACTTCTTGTACTCTTCTGTGTAT---AAGAACAATTCGTCCTCATCTTC 1035  
 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgGluValGluIlePhe 370  
 1036 TTGTTGTCTTCTTGTGCTCATCTTTGACCGGTGCGCTGTTGTTGTTGAAACCATTTTC 1095  
 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 1096 TTGGGTTCTCGTCCATGATGCCAGGTACCCCGTGTGTCGACGTTTGCACACGTT 1155  
 391 TyrTTPGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 1156 TACTGGCAATGCGTCCATTTGTTGGAATTTGTTGGTAAACACGCGCCCAATGTCATAC 1215  
 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 1216 GGTGTTTGTGAAACCCACTGTCCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1266  
 431 Leu-----Asn 432  
 1267 GCGGTGTTTGTGCGCGTGAACCAAGGTTCTGTCGCGCCCGCCAGAGAAGAAGAT 1326  
 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTTPGlnValTyr 452  
 1327 ACCGATCACCAGTCGTCGTTGGTTCAATTTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1386  
 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTTPGlyThrArgHis 472  
 1387 GGTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1446  
 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492  
 1447 AACGAACGTCGTTCTTCGCTAACCAACCAAAATTCATTTCTTTGGGTAAACACGCGCAA 1506  
 493 LeuSerLeuGlnGluLeuMetTTPLysMetLysValGluAspCysHisTTPLeuArgSer 512  
 1507 TTGTTTGTGCAAGAAATTGACCTGGAAATGCTGTTGCTGATGTCGTCGTCGTCGTCGTCG 1566

513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1567 TCTCCAGGTGTTGGTGTTCAGCGCGCAACCGGTTGCGTGAAGAAATTTGGCC 1626  
533 ThrPheLeuPheTrpLeuMetAspThrTyValValGlnLeuLeuArgSerPhePheTy 552  
1627 AATCTTGGCACTGGTGTGATCTGTCTGTTAGCTTGTGAATGTTGGTCTTCTTCTAC 1686  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyArgLysSerValTrpSer 572  
1687 GTTACCGAAACCACTTCCAAAACCGGTTGTTCTTCTACCGTAAATCTGTTGGTCT 1746  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1747 AATTCGAATCTATGTTGATTCGTCAACACTTGAACGTTCAATTCGGTGAATGCT 1806  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1807 GAAGCCGAAGTTCGTCAACACCGTGAAGCGCGTCCAGCCTTGTGACCTCTCGTTGGGT 1866  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTySerMetGlyThr 632  
1867 TCAATTCAAAACAGATGTTGCGTCCAAATGTTAAATGATGATGATGTTGGTSCC 1926  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1927 CGTACCTTCGTCGTGAAACCGTGGCGCAAGCTTTCACCTCTCGTGTAAAGCCTTGTTC 1986  
653 SerMetLeuAsnTyGluArgThrLysHisProHisLeuMetClySerSerValLeuGly 672  
1987 TCTGTTTGAACACTACGACGTCGCCGTCGTCAGGTTGTTGGGTGCCCTCTGTTGGGT 2046  
673 MetAsnAspIleTyArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2047 TTGGATGATATTCACCGTCGTCGCGTACCTCGTTTGGGTTCGTCGCCAAGATCCA 2106  
693 ThrProArgMetTyPheValLysAlaAspValThrGlyAlaTyAspAlaIleProGln 712  
2107 CCACGAAATGTTACTTCGTTAAAGTTGATGTTACCGGTGCCCTACCATACCATCCACAA 2166  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyCysIle 732  
2167 GATCGTTTACCGAAGTTATGCTCTATTATTAAAC---CCACAAACACCTACTGTGTT 2223  
733 ArgGlnTyAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2224 COTCGTTACCGCGTGTTCAAAACCGCGCCACCGTCACGTCGTTAAAGCCTTCAATCT 2283  
753 GlnValThrLeuSerAspLeuGlnProTyMetGlyGlnPheLeuLysHisLeuGln 772  
2284 CACGTTTCTACCTTGACCGATTCACACCATACATGCGTCAATTCGTTGCCCACTTGCAA 2343  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2344 GAAACC-----TCTCCATTGCGTGATGCGCGTGTGTTATTGAACAACTCTTCTTTGAAC 2397  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2398 GAAGCCTCTCTGTTGTTGTTGATGTTTCTTGGCTTCATGTTGTCACCGCGGTTCGT 2457  
813 IleGlyAspArgCysTyThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2458 ATTCTGGTGAATCTTACGTTCAATGTCAAGGTATTCACAGGTTCTATTTGTCTACC 2517  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2518 TTGTTGTGTTCTTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2577  
853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2578 GATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 927  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyGlyCysMetIleAsn 892

2638 GCCAAAACCTTCTTGGCGTACCTTGTTCGGTGTTCAGAAATACGGTTCGTGTGTTAAC 2697  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2698 TTGCGTAAACCGTGTGTTAACTTCCAGTTGAAGATGAAGCCTTGGGTGTACCGCCTTC 2757  
913 TyGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2758 GTTCAATGCGCAGCCACCGTGTGTTCCCATGGTGTGTTGTTGTTGTTGTTGTTGTTGTT 2817  
933 LeuGluValPheCysAspTySerGlyTyAlaGlnThrSerIleLysThrSerLeuThr 952  
2818 TTGGAAGTTCAATCTGATTAATCTTCTTACCGCGGTACCTCTATTCGTGCTCTCTGACC 2877  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2878 TTCAACCGTGGTTTCAAGCGCGTCTGAACATGCTGCTGAATGTTGTCGGTGTTCGGT 2937  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2938 TTGAATGTCACCTCTTGTCTTGGATTGCAAGTTAACTCTTTCGCAACCGTTGTGACC 2997  
993 AsnIleTyLysIlePheLeuLeuGlnAlaTyArgPheHisAlaCysValIleGlnLeu 1012  
2998 AACATTTACAAATTTGTTGTGCAAGCCTACCGTTTCCACGCTGTGTTTTCGCAATG 3057  
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3058 CCATTCACCAACAAAGTTGGGAAACCAACCTTCTTCTTGGTGTATTATTCGTGATACC 3117  
1033 AlaSerCysCysTyAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3118 GCCTCTTGTGTACTCTATTGTAAGCAAAACCGCGTATGCTTTCGGTGCCTCAA 3177  
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyTrpGlnAlaPheLeu 1069  
3178 GGTGCGCGCGTCCATTCGCACTCAAGCGGTCAATGTTGTGTGTCACCAAGCCTTCTTG 3237  
1070 LeuLysLeuAlaAlaHisSerValIleTyLysCysLeuLeuGlyProLeuArgThrAla 1089  
3238 TTGAAATTCACCGCTCACCGTGTATCTACCTACGTTCCATGTTGGGTCTTTCGCGTACCGCC 3297  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3298 CAACCAATGTCGTGAATTCGACAGTACCACTTGACCGCCTTGAGCGCGCCGCC 3357  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3358 AACCCAGCCTTGCCATCTGATTTTCAAAACCAATTTGGAT 3396

RESULT 6  
US-08-974-549A-721  
; Sequence 721, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

```
NAME/KEY: -  
LOCATION: 1..3451 /note= "NRTT sequence employing codon  
OTHER INFORMATION: distribution preferentially used by  
OTHER INFORMATION: highly expressed genes in E. coli  
OTHER INFORMATION: containing SacI and XhoI sites"  
;08-974-549A-721
```

	0	Length:	3451
ignment Scores:		Matches:	719
ed. No.:	3505.00	Conservative:	122
ore:	Percent Similarity:	Mismatches:	260
	72.94%	Indels:	52
	ist Local Similarity:	Gaps:	13
	ery Match:		
	59.40%		
	3:		
	3		

;-09-042-460-2 (1-1122) x US-08-974-549A-721 (1-3451)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20

28 ATGCGCGTCTCGGCTGGCCGCTGCTGTTCTGTTCCCTGCTGCGTTCCCATCTATCCGGAA 87

21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40

88 GTTCTGCGGCTGGCTACCTTCGTTCTGCTCTGCGCGCCGACGGGTGGCGTCTGTTCTCAG 147

41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60

148 CGTGGTAGCCGGCTCTTCCGTGCTCTGGTTGTCTCAGTGCCCTGGTTGGGTTCCGTGG 207

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80

208 GACGCTCGTCCGCGCGGCTGCTCGTCTCTTCCGTACAGTTTCTGCTCTGAAAGACTG 267

81 ValAlaArgValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheCly 100

268 GTTCTCGTGTCTTCGACGGCTCTGTCGCAACGTGGTGTAAAAAGCTTCTGGCTTTGGT 327

101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120

328 TTCGCTCTGCTGGACGGTGTCTGTTGGTCTCCGCGGAAGCATTCACCACTCCGTTCTG 387

121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140

388 TCTTACTGCGCAACACCGTTACGACGCTCTGCGTGTTCGGTGTCTGGGCTCTGCTG 447

141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160

448 CTGGCTGTGTTGGTAGACGTTCTGGTTTACCTGTGCTGCTGCTGCTGCTGCTGCTG 507

161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180

508 CTGGTTGCTCGTCTCGCTTACCAAGTTTGTGCTCCGCGCTGTACAGCTGGGTGCT 567

181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200

568 GCTACCACAGCTCGTCCGCGCGCGACGCTTCC---GGTCCGCTGCTGCTGCTGCT--- 621

201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220

622 -----TCGGAACGCTGTGGAACCACTCCGTTTCGTGAAGCTGGT 660

221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240

661 GTTCGCTGGGTCGTGCGGTCGCGGTGCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCT 720

241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259

721 CTGCGCTGCGGAACGCTCGCGTCTGGTGTGCTGCTCCGGAACCGGAACGTTACCCGGTT 780

260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272

781 GGTGAGGTTCTGGGCTCACCCGGGTGCTACCCGTGTGCTGCTGCGACCGTGTCTTCG 840

273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292

841 GTTGTTCCTCCCGGCTCGT-----CCGGCTGAAGAACTACCTCCCTCGGA 885

293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311

886 GGTGCTCTGTCCGGCACCCGCTCACTCCACCCGCTCGTTGGTGTGCTGACCAACGCTGGT 945

312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330

946 CGCGCGTCCACTCCCGTCCGCGCGCTGCGTGGGACACCCCGTGTCCGCGGTTTACCGT 1005

331 GluThrArgHisPheLeuTyrSerArgLysGlyAspGlyGlnGluArgAsnProSerPhe 350

1006 GAAACCAACACTCTCTGTACTCTCCGGTGAC---AAAGAACAGCTGCGTCCGCTCTTC 1062

351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370

1063 CTGCTGTCTCCCTGCGCTCGGTCCTGACCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1122

```
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
1123 CTGGGTTCCCGTCGTCGATCGCGGACCCCGCGTCGTCGCGCGTCGTCGCGAGGCT 1182
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGlnCysGlnTyr 410
1183 TACTGGCAGATGCGTCGCTGTTCTCGGAACCTGCTGGCAACACCGCTCAGTGGCCGAT 1242
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
1243 GGTGTTCTGCTGAAACCAACCTGCTCCCGCTCGCTGCT-----GTACCCTGGCT 1293
431 Leu-----Asn 432
1294 GCTGGTGTTCGCTGCGTGAACCAACCGAGGTTCCGTCGCTCCGGAAGAAGAT 1353
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
1354 ACCGACCGCGTCGTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
1414 GGTTCGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheLeuLeuLysLysLysLysLys 492
1474 AACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
1534 CTGTCCTGCGAGAACTGACCTGGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
513 SerProGlyLysAspArgValProAlaAlaGlnHisArgLeuArgGluArgLeuAla 532
1594 TCTCCGGGTGTTGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1653
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552
1654 AAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1713
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
1714 GTTACCGAAACACCTCTCCGAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1773
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1774 AAATGTCAGTCCATCGGTATCCGTCAGCACCTGAAACGCTGTTTCCAGTGGTGAACCTCC 1833
593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProLleCysArgLeuArg 612
1834 GAAGCTGAAGTTCGTCAGCACCGTGAAGCTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1893
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1894 TTTCATCCCGAAACCGAGCGGTCTCGTCCGATCTGAACATGACCTACGTTGTTGGTCT 1953
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1954 CGTACCTCTCCGTCGTGAACAAACCTGCTGAGCGCTGACCTCCCGCTGTTAAAGCTCTGTT 2013
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
2014 TCGTTCCTGACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2074 CTGACGACATCCACCGTCTGGCGTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2133
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2134 CCGCCGGAACCTGCTACTTCTGTTAAAGTTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2193
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2194 GACCGTCTGACGAGGTATCCGTTCCATCATCAAA-----CCGCAACACCTACTGCTT 2250
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752
2251 CGTCGTTACGCTGTTGTTTCAGAAAGCTGCTCAGCGTCACGTTGTTAAAGCATTCAAATCC 2310
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisIleGln 772
2311 CAGGTTTCCACCTGACCGACCTGACGCGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2370
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2371 GAAACC-----TCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2424
793 GluSerSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
2425 GAAGCTTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2484
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2485 ATCCGTTGTAATCTTACGTTTCCGTCAGCGGATCCGCGAGGTTCCATCTCTGCTCCACC 2544
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2545 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2604
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872
2605 GACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2664
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2665 GCTAAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2724
893 LeuGlnLysThrValValAsnPheProGluProGlyThrLeuGlyGlyAlaAlaPro 912
2725 CTGCGAAACCGCTGTTTAACTTCCCGGTGAAGACGAGCTCTCGGTGGGACCGCTTTC 2784
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932
2785 GTTCAGATGCGGCTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2844
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2845 CTGGAAGTTCAGTCCGACCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2904
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2905 TTCAACCGTGGTTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2964
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2965 CTGAAATGCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3024
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3025 AACATCTACAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3084
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3085 CCGTTCACCGAGGTTTGGAAAAACCGAGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3144
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
3145 GCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3204
1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
3205 GGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3264
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
```

3265 CTGAACTGACCGTCCCGTGTACCTACGTTCCGCTGCTGGGTTCCCTGCGTACCGCT 3324  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3325 CAGACCCAGCTGTCCCGTAACTGCGGGTACACCTGACCGCTCTGGAAGCTGCTGCT 3384  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3385 AACCCGCTGCTGCGTCCGACTTCAAAACCATCTCGAC 3423

## SULT 7

-08-851-843A-224

Sequence 224, Application US/08851843A

Patent No. 6093809

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: NO. 6093809el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:  
SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTERT"

OTHER INFORMATION: /note= "human telomerase reverse

OTHER INFORMATION: transcriptase (hTERT) catalytic protein

## OTHER INFORMATION: component "

US-08-851-843A-224

Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 3 Gaps: 13

US-09-042-460-2 (1-1122) x US-08-851-843A-224 (1-4015)

QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
DB 56 ATGCCGGCGCTCCCGCTGCGAGCGGTGCTCTCCCTGCTGCGCAGCCTACTCCGGAG 115  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
DB 116 GTGCTGCGCTGCGCACGTTGCTGCGCGCTGCGCGCCCGCAGGCTGCGGCTGCTGCG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
DB 176 CCGGGGACCCCGCGCTTTCGGCGCTGTGCGCCAGTGCCTGCTGCTGCTGCTGCTGCTG 235  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
DB 236 GACGACGCGCGCGCGCTTCTCCCTTCCGCGAGGTGCTCTGCTGCTGCTGCTGCTGCTG 295  
QY 81 ValAlaArgValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
DB 296 GTGGCCCGAGTGTCTGACAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
DB 356 TTCGCGCTGCTGACG 415  
QY 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
DB 416 AGCTACCTGCCCAACACCGGTGACCGACTGCGGGGCGCGCGCGCGCGCGCGCGCGCGCT 475  
QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
DB 476 CTGCGCGCGCTGCGCGAGCAGCGTGTGCTTCACTGCTGCGCAGCTGCGCGCTCTTTGTG 535  
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
DB 536 CTGCTGCTCCAGCTGCGCTTACCAGGTGTGCGCGCGCGCGCGCGCGCGCGCTGTACCA 595  
QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
DB 596 GCCACTCAGCG 649  
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
DB 650 -----TGCGACGCGCGCTGGAACCACTAGCTCAGCGGAGCGCGCG 688  
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
DB 689 GTCCCTTGGCTTCCAGCCCGCGGTGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGAGT 748  
QY 241 ValProSerAlaLysLeuAlaArgCysTyrProValProArgValGluGluPro--- 259  
DB 749 CTGCGCTTGCACAGAGCGGT 808  
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
DB 809 GGCAGGGGTCTCTGGCGGT 869  
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
DB 869 GTGGTCTCAGCTGCCAGA-----CCCCCGCGAAGAACCCCTCTTTGGAG 913

```
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
914 GGTGGCTCTCTGGCAGCGGCCACTCCACCCATCCGTGGCGCGCCAGCACACGCGGCG 973
312 SerThrSerLeuLeuSerProProAspGlnAsnAlaPheGlnLeuArgPro---Phe1le 330
974 CCCCATCCACATCGCGGCCACACAGTCCTCGGACACGCGCTTGTCCCGGTGTACGCC 1033
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
1034 GAGACCAAGCACTCTCTACTCTCTCAGGGGAC---AAGGAGCAGCTGGCGCCCTCTTC 1090
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGlu1le1lePhe 370
1091 CTACTCAGCTCTCTGAGGCCAGCGCTCAGTGGCGCTCGAGGCTCTGGAGACCATCTTT 1150
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
1151 CTGGGTTCCAGGCCCTGGATCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCGC 1210
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
1211 TACTGGCAATGCGGCCCTCTTCTGGAGCTGCTTGGGAACACAGCGCGAGTCCCTTAC 1270
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTCG-----GTACCCCGCAGCA 1321
431 Leu-----Asn 432
1322 GCGGTGTCTGTGCCGGGAGAACCCAGGGCTCTGTGGCGGCCCGGAGGAGGAGC 1381
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
1382 ACAGACCCCGCTGGCTGGTGCAGCTGCTCGCCAGCACAGCAGCCCTCGCGAGGTGAC 1441
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
1442 GCGTCTGTCGGCGCTGCTGCGCGGTGTGTGCCCGCCAGCGCTCTGGGGCTCCAGCGAC 1501
473 AsnGluArgArgPhePheAsnLeuLysLysPhe1leSerLeuGlyLysTyrGlyLys 492
1502 AACGAACGCGCTTCTCAGAACACCAAGAGTTTCATCTCCCTGGGAGAGATGCCAAG 1561
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
1562 CTCTGCTCAGAGCTGACGTGGAAGATGAGCGTGGGACTCGCTTGGCTGGCGAGG 1621
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgG1leLeuAla 532
1622 AGCCAGGGGTGGCTGTGTTCCGCGCCGAGACACCGCTCTGCGGTGAGGAGATCCTGGCC 1681
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
1682 AAGTTCCTGCACTGGCTGATGATGTGTACGTCTGTGAGCTGCTCAGTCTTTCTTTAT 1741
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
1742 GTCACGAGACCAACGCTTCCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
1802 AAGTTCGAAGCATTTGATCAGACAGCATTTGAAGAGGTGCAGCTCGGGAGCTGTGG 1861
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetPro1leCysArgLeuArg 612
1862 GAAGCAGAGGTGAGGATCGGAAGCAGGCGCCCTGCTGACGTCCAGACTCCGC 1921
613 PheIleProLysProAsnGlyLeuArgPro1leValAsnMetSerTyrSerMetGlyThr 632
1922 TTCATCCCAAGCCTGACGGGTGGCGGATTTGACATGGAGTGGACTAGCTGGGGAGCC 1981
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnA-gLeuLysThrLeuPhe 652
```

---

```
1982 AGAACCTTCGCGAGAGAAAGAGGCGCGCTCTCACCTGAGGGTGAAGGCACTGTTC 2041
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
2042 AGCGTCTCAACTACGAGCGGCGCGCGCGCTCTCTGGGCGCTCTGTGTGTGGGC 2101
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValArgValArgAlaLeuAspGln 692
2102 CTGAGCATATCCACAGGCGCTGGCGCACCTTCGTGCTGCGGTGCGGGCCAGGACCGC 2161
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAla1leProGln 712
2162 CCGCTGAGCTGACTTTCTCAAGGTGGATGTGACGGGGGTACGACCATCCCCAG 2221
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCys1le 732
2222 GACAGGTCTACGAGGTCTATCCAGCATCATCAAA---CCCCAGAACACGTACTGCGTG 2278
733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2279 CGTCGTATGCGCTGTCCAGAGGCGCGCCATGGGCACGTCCGCAAGGCTTCAAGAGC 2338
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2339 CAGCTCTCTACCTTGACAGCCTCCAGCGGTATACGCGACAGTTCGTGGCTCACCCTGAG 2399
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792
2399 GAGACC-----AGCCCGCTGAGGATGCGCTCATCGACGACAGCTCCTCCCTGAAT 2452
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
2453 GAGCCAGCAGTGGCTCTTCGACGTCTTCACGCTTCATGTGCCACACCGCGCTGGC 2512
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2513 ATCAGGGCAGTCTTACGTCCAGTCCAGGGAGTCCCGCAGGGCTCCATCTCTCCAGC 2572
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2573 CTGCTCTGACGCTGTCTACGCGCAGCATGGAGAACAGCTGTGTGCGGGGATTCGGCGG 2632
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2633 GACGGGTGCTCTGCGCTTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2692
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2693 GCGAAACCTTCCTCAGGACCCCTGGTCCGAGGTGTCCCTGAGTATGGCTGCGTGTGAAC 2752
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2753 TTGCGGAGACAGTGGTGAACCTTCCTCTAGAGACGAGCGCCCTGGGTGGCAGCGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATGCGCGCCACCGCTATTCCTCTGTGTGGCTGTCTGTGATACCGCGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGAGGTGACAGCGCACTACTCCAGCTATGCCCGGACCTCCATCAGACCGCAGTCTCAC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2933 TTCACCGCGCTTCAAGGTGGAGGAACATGCTGCGCAACTCTTTGGGTCTTGGCG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCys1le 992
2993 CTGAAGTGTACAGCCCTGTTCTGGATTTGAGTTGACGTTGAACAGCCCTCCAGACGGTGTGACC 3052
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
```

3053 AACATCTACAGATCTCTCTGCTGAGGGGTACAGGTTTACGCATGTGTCTGACGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCATCAGCAAGTTTGGGAAGAACCCCAATTTTCTCGCGCTCATCTCTGACACG 3172  
1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCTCTGCTACTCTCATCTCTGAAGCAGAGAGCGAGGATGCTCGGGGGCCAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3233 GCGCGCGCGCGCTCTGCGCTCCCGAGCGGTGAGTGGCTGTGTCACCAAGCATCTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTGACGACCGTGTACCTACGTGCGACCTCTCTGGGTCTACTCAGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGCGCTGAGTGGGAAGCTCCCGGGGAGCGCTGACTGCTGCGCTGGAGCGCGACGC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGCGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3451

## SULT 8

-08-974-549A-1

Sequence 1, Application US/08974549A

Patent No. 6166178

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "hTERT"  
OTHER INFORMATION: /note= "human telomerase reverse transcriptase (hTERT) catalytic protein component"  
US-08-974-549A-1  
Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: 3 Gaps: 13  
US-09-042-460-2 (1-1122) x US-08-974-549A-1 (1-4015)  
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCCGCGCGCTCCCGCTGCCGAGCGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
QY 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCGCTGGCCACGTTCTGCGCGCGCTGGGGCCCCCAGGGCTGCGCGGTGTCAG 175  
QY 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
Db 176 CCGGGGACCGCGCGCTTCCGCGCGCTGGTGGCCAGTCCCTGGTGGTGGTGGTGGTGG 235  
QY 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GACGACG 295  
QY 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGGCCCGAGTGTCTGACAGAGGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 355  
QY 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
Db 356 TTCGCGCTGCTGAGCGGGGCCCGCGGGGGCCCCCGCGAGGCTTTCACCCAGCGGTGCG 415  
QY 121 SerTyrLeuProAsnThrValIleGlnThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
Db 416 AGCTACCTGCCACACCGGTGACCGCACTGGGGGGGCGCGGGGCGTGGGGGCTGCTG 475

```
Y 141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
b 476 CTGCCCGCGTGGCGAGCGAGTGTCTGTTTCACTGTCTGGCAGCTCGCGCTTTTGTG 535
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla 180
b 536 CTGGTGGCTCCAGCTGGCTTACAGGTGTGGCGCGCGCTGTACCACTCGGCGCT 595
Y 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
b 596 GCCACTCAGCGCGCGCGCGCCAGCACACTAGT---GGACCCCGAAGCGCTCTGGGA--- 649
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
b 650 -----TGCGAAGCGCGCTGGAACTACCTACGTCAGGGAGCGCGG 688
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
b 689 GTCCCCCTGGGCTGCGAGCGCGCGTGGAGAGCGCGCGGCGAGTGGCCAGCGCAAGT 748
Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
b 749 CTGCCCTTCCCAAGAGCGCGAGCGTGGCGCTGCCCTTGAGCCGGAGCGGACGCCGTT 808
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
b 809 GGCGAGGGTCTTGGGCGCCACCGCGCGAGGACGCGTGGACCGAGTACCGTGGTCTCTGT 868
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
b 869 GTGGTGTACCTGCCAGA-----CCGCGCGAAGAGCGACACTCTTTGGAG 913
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
b 914 GTGGCGCTCTTGGCAGCGCGCACTCCACCCACTCCGTGGCGCGCCAGCACCGCGGC 973
Y 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
b 974 CCCCACATCCACATCGCGCGCACACGTCCTCGGACACGCCCTTCTCCCGCGGTATAGCC 1033
Y 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
b 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGC---AAGGAGAGCTGCGCGCTCTCTTC 1090
Y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
b 1091 CTACTCAGCTCTCTAGGCGCGAGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTTT 1150
Y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
b 1151 CTGGGTTCAGGCGCTGGATCGCAGGACTCCCGCAGGTTGCCCGCGCTGCCCGCAGCG 1210
Y 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
b 1211 TACTGGCAATCGCGCGCTGTTCTTGAGCTGCTTGGAAACACCGCAGTGCCTCTAC 1270
Y 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
b 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGGAGTGGC-----GTACCCCGCAGCA 1321
Y 431 Leu-----Asn 432
b 1322 CGCGGTGTCTGCCCGGAGAGCCCGAGGCTCTGTGGCGCGCCCGAGGAGGAGGAC 1381
Y 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
b 1382 ACAGACCCCGCTGCGTGGTGGAGTGTCTCCCGCAGCACAGCAGCCCGTGGCAGGTATC 1441
Y 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472
b 1442 GGCTTGTGGCGGCTCTGCTGCGCGGTGTGGTGGCGCGGCTCTGGGCTCCAGGCAC 1501
```

```
QY 473 AsnGluArgArgPhePheLeuLysAsnLeuLysLysPheLeuLysTyrGlyLys 492
Db 1502 AACAGACCGCGCTTCTCAGAAACACACAGAGATTCTCTCTGGGAGCATGCCAAG 1561
QY 493 LeuSerLeuGlnLeuLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512
Db 1562 CTCTCGCTCAGGAGCTGACGTGGAAAGATGAGCGTCCGGGACTCGCTTGGCTGGCGAG 1621
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
Db 1622 AGCCACGGGTGGCTGTGTCTCCGGCGCAGACACCTCTTCTTACCGAAGAGTGTCTGGAG 1681
QY 533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
Db 1682 AAGTTCTGCTGACTGGCTGATGAGTGTGTACGTCTGTGAGTGTCTCAGTCTTCTTTAT 1741
QY 553 IleThrGlnSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrPhe 572
Db 1742 GTCCGAGAGACCGTTCATAAAGAAACAGAGCTCTTTTCTACCGAAGAGTGTCTGGAG 1801
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
Db 1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTCAAGAGGTTGAGTGGCGGAGCTGTCTG 1861
QY 593 GlnGlnGluValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgLeuArg 612
Db 1862 GAACGAGAGTCAAGCAGCATCGGAGAGCGCGCCCTGTGTGACGTCCAGACTCCCG 1921
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
Db 1922 TTATCTCCAGCTTACGCGCTGCGGCTGCGGCGATTTGTAAATGGACTACCTCGTGGAG 1981
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
Db 1982 AGAACGTTCCGACAGAAAGAGGCGCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTC 2041
QY 653 SerMetLeuLeuTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
Db 2042 AGCTGTCTTAACACAGAGCGCGCGCGCGCGCTCTGGGCGCTCTGTGTGGGCG 2101
QY 673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692
Db 2102 CTGGACGATATCCACAGGCGCTGGCGCACCTCTGTCTGTGGCGCGCCAGGACCGG 2161
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
Db 2162 CCGCTGAGCTGTACTTTGTCAAGGTGATGTACCGGCGCGGTACGACACCATCCCGCAG 2221
QY 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
Db 2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCGAAGACAGTACTGCGTG 2278
QY 733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
Db 2279 CGTCGATGCGGTGTCTCAGAAAGCGCGCCATCGGCGCGTCCCGAAGCGCTTCAAGAG 2338
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
Db 2339 CAGCTCTCTACTTTCAGACACTCCAGCGCTACATCGGACAGTTCGTGCTCACCTGCGAG 2398
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792
Db 2399 GAGACCC---AGCCGCTGAGGATGCGCTGTCTATCGAGCAGAGCTCTCTCCCTGAAT 2452
QY 793 GluSerSerSerLeuPheAspPhePheHisPheLeuArgHisSerValValLys 812
Db 2453 GAGCGCACAGTGGCTCTTCGAGCTTCTTACGCTTCTATGTCACACCGCGCGTGGCG 2512
QY 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
Db 2513 ATCAGGGGCAAGTCTCAGTCCAGTCCAGGGGATCCCGAGGGCTCCATCTCTCCAG 2572
QY 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
```

```
2573 CTGCTCTGAGCTGTGCTACGGCAGCATGGAGAAAGCTGTGTCGGGGATTCGGCGG 2632
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872
2633 GACGGGCTGCTCTGCTGTGCTGTGATGATTTCTGTTGGTGACACCTCACCTCACCCAC 2692
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleasn 892
2693 GCGAAACCTTCTCAGGACCTGTGTCAGGTGCTCCCTGAGTATGGCTGGTGGTGAAC 2752
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyValAlaAlaPro 912
2753 TTGCGGAGACAGTGTGTGACTTCCCTGTAGAGACGAGGCGCTGCTGGTGGCAGCGCTTTT 2812
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2813 GTTCAGATCGGGCCCGCCAGCGCTATTCCCTGTGTCGGCTGCTGCTGATACCGGACC 2872
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2873 CTGGAGTGCAGGCGACATCTCCAGCTATGCGGACCTCCATCAGACCGACGCTCACC 2932
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2933 TTCACCGCGGCTTCAAGCTCGGAGGAAATGCTGCGCAAACTCTTTGGGCTCTTGGG 2992
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
2993 CTGAAGTGTACAGCTGTTCTTGGATTTCAGGTGAACAGCTCCAGCGCTCCAGCGGTGACC 3052
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
3053 AACATCTACAAGATCTCTCTGCTGCGAGGCGTACAGCTTCACGCACTGTGCTGAGCTC 3112
1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
3113 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTCGCGCTCATCTCTGACACG 3172
1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaser 1052
3173 GCCTCCCTCTGCTACTCCATCTCTGAAGCCAGACGAGGATGCTCGTGGGGGCCAAG 3232
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069
3233 GGGCGCGCGGCGCTCTGCGCTCCGAGGCGGTGACGTGCTGTGCCACCAAGCATTCCTG 3292
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
3293 CTCAAGCTGACTCGACACCGTGTACCTACGTGCGCACTCTCTGGGTCTACTCAGGACGCC 3352
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109
3353 CAGACGCGAGTGTAGTCCGAAGCTCCCGGGGACGACGCTGACTGCTCGCGGCGGCGGCG 3412
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
3413 AACCGGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3451
```

SULT 9

-08-854-050-224

Sequence 224, Application US/08854050

Patent No. 6261836

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Haxley, Calvin

APPLICANT: Andrews, William H.

TITLE OF INVENTION: No. 6261836gel Telomerase

NUMBER OF SEQUENCES: 225

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002930US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 224:
SEQUENCE CHARACTERISTICS:
LENGTH: 4015 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 56..3454
OTHER INFORMATION: /product= "hTERT"
OTHER INFORMATION: /note= "human telomerase reverse
OTHER INFORMATION: transcriptase (hTERT) catalytic protein
component"
US-08-854-050-224

```

## Alignment Scores:

Pred. No.:	0	Length:	4015
Score:	3505.00	Matches:	719
Percent Similarity:	72.94%	Conservative:	122
Best Local Similarity:	62.36%	Mismatches:	260
Query Match:	59.40%	Indels:	52
DB:	3	Gaps:	13

US-09-042-460-2 (1-1122) x US-08-854-050-224 (1-4015)

Qy	1	MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu	20
Db	56	ATGCGCGCGCTCCCGCTGCGAGCGCGTCCCTGCTGCGCAGCCACTACCGCGAG	115
Qy	21	ValTyrProLeuAlaThrPheValArgGluGlyProGluGlyArgArgLeuValGln	40
Db	116	GTGCTGCGCGCTGGCCACGTTCTGTCGCGCGCTTGGGGCCCCCAGAGGCTGGCGGCTG	175



41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
176 CGCGGGACCGCGCGGCTTTCGCGCGGCTGGTGGCCAGTGGCTGGTGGCTGGCTGG 235  
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
236 GACGACGCGCGCGCGCGCGCGCTTCTCCGCGAGGTGTCTGCGCTGAAGAGGCTG 295  
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
296 GTGGCCCGAGTGTGACAGAGCTGTGCGAGCGCGCGCGAAGAACGTGTGCGCTTCG 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerValArg 120  
356 TTCGCGTGTGACG 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeu 140  
416 AGCTACCTGCCACACAGTACCGACGACCTGCGGGGAGCGGGCGGTGGTGGT 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTGCG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTGGTGGCTCCAGCTGGCTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCT 595  
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCG 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGCACGCGCGCTGGACCATAGCTCAGGAGCGCGCG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240  
689 GTCCCGCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
241 ValProSerAlaLysLeuAlaArgCysTyrProValProArgValGluGluPro 259  
749 CTGCGGTGCGCACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
809 GGGCAGGCGTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTACCTGCGCAGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysLysHisLysProSer 311  
914 GTGCGCTCTGTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro-----PheIle 330  
974 CCCCCATCCACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTCTACTCTCAGCGC-----AAGAGCAGCTGCGCGCTCTTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410

1211 TACTGCAAAATCGCGCGCGCTGTTCTGGAGCTGCTGGAAACACGCGCGAGTCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGCTGG-----GTCAACCCAGCA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTCGCGCGGAGAGCCCGAGGCTCTGTGGCGCGCGCGCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCGCTGGTGCAGCTGCTCGCGCAGCAGCAGCCCTGCGAGTGTAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGCGCGCGCTGCTGCGCGCGGTGGTGGCCCGCGCGCGCTGCGGGCTCCAGGAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AACCAACGCGCTTCTCAGGAACACCAAGAGTTCACTCCCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGCGAGAGTGCAGTGCAGAGATGAGCGTGCAGGCTGCGCTTGGCTGCGCAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGGCGTGGCTGTGTTCCGCGCGCAGCAGCAGCTCTGCGTGCAGAGATCTCTGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
1682 AAGTCTCTGCACTGGCTGATGATGTGTACGTGTCGTCGTCGAGCTCTCTCTTTAT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTACGAGACCGAGTTCAAAGAACAGCGCTCTTTCTCCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgLysLeuSer 592  
1802 AAGTTCGACAGCATTCGAATCAGACAGCATTTGAAGAGGTGCAGTGCAGGAGCTGCG 1861  
593 GlnGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGTCCAGGAGCATCGGAAGCAGCGCGCGCGCGCTGCTGACGTCCAGATCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCGAGCTCAGCGGCTGCGCGCGATGTTGTGAACATGAGTACGTCTGCGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGCGCGCTCTCACCTCGAGGGTGAAGGCATCTGTTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCACTACAGCGCGCGCGCGCGCGCTCTGCGCGCTCTGCGCGCTCTGTTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACCATATCCAGCGCGCTGCGCACCTTCTGCTGCTGCTGCGCGCGCGCGAGCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGAGTGTACTTGTCAAGGTGGATGTACGCGCGCGTACGACCATCTCCCGAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCAGCATCATCAAA-----CCCCAGAACACGTACTCGGTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752

2279 CGTCGGTATGCCGTGCTCCAGAGGCGCGCCATGGGACGCTCCGACGCTTCAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGGTCTCTACCTTGCACAGACCTCCAGCGGTATACGCGACGTTGCTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuSerAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGAC-----AGCCGCTGAGGGATGCGCTGTCATGACGACGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCCAGCAGTGGCTCTTCGACGCTTCTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTTACGCTCAGTCCAGGGGATCCCGACGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGlnAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCAGCTGTGCTACGCGGACATGGAGAACAGCTGTTGCGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTGCTGCTTGTGGTGGATGATTTCTTGTGGTGCACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTTGTTCGAGGTGTCCTGAGTATGCTGCTGGTGGTGCAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGTGTGACTTCTCTGTAGAGACGAGCGCTCGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGGCGCCACGCGCTTATTCCTGCTGGTGGCTGCTGCTGATACCCGAC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGCAGAGCAGCTACTCCAGCTATGCCGAGCTCCATCAGACGACGCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGGCTTCAAGGTGGAGAGACATGCTGCGCAAACTTTTGGGGTCTTGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCGCTGTTCTGGATTTCAGGTGAACAGCGCTCCAGCGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCTGCGAGCGGTACAGGTTTCACGCGATGTGTGCTGCG 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCATCAGCAAGTTTGAAGAACCCACATTTTCTGCGGTCTATCTCTGCACG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCTCTGAAAGCAAGACGAGCGGATGCTGCTGGGGCCCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GGGCGCGCGCGCTCTGCGCTCCGAGGCGGTGCGAGTGGCTGTGCACAGCATTCCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAAGCTGACTGCACACCGGTGTCACTACGTGCCACTCTCTGGGGTCACTCAGGACG 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACGAGCTGAGTGGAGAGCTCCCGGGGACGACGCTGACTGCTGCGGCGCGACCC 3412

QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3451  
RESULT 10  
US-09-430-323-224  
; Sequence 224, Application US/09430323  
; Patent No. 6309867  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. 6309867el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/430.323  
; FILING DATE: 29-Oct-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4015 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 56..3454  
; OTHER INFORMATION: /product= "hTRT"  
; /note= "human telomerase reverse  
; transcriptase (hTRT) catalytic protein  
; component"  
; SEQUENCE DESCRIPTION: SEQ ID NO: 224:  
US-09-430-323-224

Alignment Scores: 0 Length: 4015  
Pred. No.: 3505.00 Matches: 719  
Score:

Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Very Match: 59.40% Indels: 52  
3: 4 Gaps: 13

3-09-042-460-2 (1-1122) x US-09-430-323-224 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCTGCTGCGCGAGCCACTACCGCGAG 115  
21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluArgArgValGln 40  
116 GTGTCGCGCTGGCGACGTTCTGTCGCGCGCTGCGCGCGCGCGCGCGCGCGCG 175  
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
176 CGCGGGGACCGCGCGCTTCCGCGCGCTGTCGCGCGCGCGCGCGCGCGCGCGCG 235  
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyLeu 80  
236 GACGCGACG 295  
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
296 GTGGCGCGCGCTGCGAGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
356 TTCGCGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140  
416 AGCTACTGCGCCACACCGTACCGACGCTGCGAGCGCGCGCGCGCGCGCGCGCG 475  
141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
476 CTCGCGCGCGTGGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
536 CTCGTGGCTCCCGAGCTGGCGCTACAGGTGTGCGGCGCGCGCGCTGTACAGCT 595  
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGCG 649  
201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGluAlaPro 220  
650 -----TGGGACCGGCGCTGGAGCCATAGCTGACGAGGCGCGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCGCTGGCGCTCCAGCGCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCG 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro 259  
749 CTCGCGCTGCG 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr 272  
809 GGGCAGGGGTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTACCTGCCAGA-----CCCGCGGAGGAGGCGCGCGCGCGCGCG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGCGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330

974 CCCCCATCCCATCGCGGCGCACACAGTCCCTGGGACAGCGCTTGTCCCGCGGTACGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACCTTCTCTACTCTCTCAGCGCAC---AAGGAGCAGCTCGCGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGTTCACAGCGCGCTGATGCCAGGACTCCCGCAGGTTGCCCGCGCTGCCCGCGC 1210  
391 TyrTyrGlnMetArgProPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATCGCGCGCTGTTCTGGAGCTGCTTGGGAACACCGCGCGCGCGCGCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGCGACTGCCCGCTGCGAGTGGG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCCGTGCTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGCGCGCGCGCGCGAGGAGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452  
1382 ACAGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472  
1442 GGCTTCGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLeuSerLeuGlyLysTyrGlyLys 492  
1502 AACGAACCGCGCTTCTCAGGAACACCAAGAGAGTTCATCTCCCTGGGAGAGATCCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512  
1562 CTCTCGCTGCGAGGCTGACGTGGAAGATGAGCGTGGGAGCTGGCTGGTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1681  
533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1682 AAGTTCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572  
1742 GTACGAGGACCACTTTCAAAAGAACAGCGCTTTTCTACCGGAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGAAGCATTTGAATCAGACAGCTTGAAGAGGCTGACAGCTGCGGAGCTGTG 1861  
593 GlnGluValValArgHisHisGlnAspThrTyrLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGGTTCAGGAGCATCGGAGGAGCGCGCGCGCGCTGCTGCTGCTGCTGCTG 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTTCAGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGTTCCGAGAGAAAGAGCGCGCGCGCTCACCTCGAGGAGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCTCACTACGAGCGCGCGCGCGCGCGCGCTCTCTGGCGCGCTCTGTCTG 2101

QY	1033	AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer	1051
DB	3173	GCCTCCCTCTGCTACTCCATCTCTGAAGAACGACGAGGATGTCTCGGGGCCAAG	3232
QY	1053	GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu	1089
DB	3233	GGCGGCGCGCCCTCTGCCTCCGAGGCGGTGCAGTGGCTGTGCACCAAGCATTCCTG	3292
QY	1070	LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla	1089
DB	3293	CTCAAGCTGACTGCACACACCGTGTCACTACGTGCACCTCTGGGGTCACTCAGGACGCC	3352
QY	1090	GlnLysLeuLeuCysArgLysLeuProGluAlaAlaThrMetThrIleLeuLysAlaAlaAla	1109
DB	3353	CAGACGCGAGCTGAGTCCGAAGCTCCCGGGGACGACGCTGACCTGGAGCGCCGACGCC	3412
QY	1110	AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp	1122
DB	3413	AACCGGCACTGCCCTCAGACTTCAAGACCATCTGGAC	3451
RESULT 11			
US-09-572-423B-3			
; Sequence 3, Application US/09572423B			
; Patent No. 6331399			
; GENERAL INFORMATION:			
; APPLICANT: Brett P. Monia			
; APPLICANT: William A. Gaarde			
; APPLICANT: Edward Wanciewicz			
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION			
; FILE REFERENCE: ISPH-0462			
; CURRENT APPLICATION NUMBER: US/09/572,423B			
; CURRENT FILING DATE: 2000-05-16			
; NUMBER OF SEQ ID NOS: 29			
; SEQ ID NO 3			
; LENGTH: 4015			
; TYPE: DNA			
; ORGANISM: Homo sapiens-			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (56)...(3454)			
US-09-572-423B-3			
Alignment Scores:			
Pred. No.: 0 Length: 4015			
Score: 3505.00 Matches: 719			
Percent Similarity: 72.94% Conservative: 122			
Best Local Similarity: 62.36% Mismatches: 260			
Query Match: 59.40% Indels: 52			
DB: 4 Gaps: 13			
US-09-042-460-2 (1-1122) x US-09-572-423B-3 (1-4015)			
QY	1	MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu	20
DB	56	ATGCGCGCGCTCCCGCTGCCGAGCGGTGGCTCCCTGCTGGCAGCACCTACCGGAG	115
QY	21	ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln	40
DB	116	GTGCTGCGCTGGCGACGCTGCTGTCGGCGCTCGGGGCCCGCGGGCTGGGTGTGCGAG	175
QY	41	ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp	60
DB	176	CCGGGGACCCGGCGGCTTTCCGGCGCTGGTGGCCAGTGTCTGGTGTGGCGCTGG	235
QY	61	GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu	80
DB	236	GACGACGCGCGCCCGCCCGCCCTCTCCGCGCAGTGTCTGCTGCTGAAGAGCTG	295
QY	81	ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly	100
DB	296	GTGGCCCGAGTCTGTCGAGAGGCTGTGCGAGCGGCGGCGAAGAACTGTCGCTTCGCG	355
QY	101	PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg	120



813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATAGGGGCAAGTCTACGTCCAGTCCAGGGGATCCCGCAGGGTCTCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuLeuPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCAGCTGTGCTACGCGCAGCATGGAGAACAACTGTTTGGGGGATTCCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTCGGTTTGGTGGAGATTTCTTGTGTGTCACACCTCACCTCACCAC 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTTGGTCCGAGGTGTCTCTGAGTATGGCTGGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGTGTGAATCTTCCCTGTAGAGACGAGGCCCTGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGGCGCCACGCGCTATTTCCCTGTGTGGGCTGTGCTGTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGCAGAGCGACTACTCCAGCTATGCCGAGCTCCATCAGAGCCAGTCTCACC 2932  
953 PheGlnSerValPheValAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTACACCGCGGCTTCAAGGCTGGAGGAGACATGCGTCGCAACTCTTTGGGGTCTTGGG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACAGCCTGTTCTTCGATTTCAGTGAACAGCTCCAGAGCTCCAGAGCTGTCAC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCGCAGGCGTACAGGTTTCACGCAATGTGTGTCGAGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuLeuPhePheLeuGlyIleIleSerSerGln 1032  
3113 CCATTTCATCAGCAAGTTTGGAGAACCCCACTTTTCTGCGGCTCATCTCTCAGACG 3172  
1033 AlaSerCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGCTACTCATCTCTGAAAGCAAGAACGAGGAGTGTGCTGGGGGCCAAG 3232  
1053 GlySer-----PheProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
3233 GGGCGCGCGGCGCTCTGCGCTCCGAGGCGGTGAGTGTGCTGTGTCACCAAGCATCTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAAGTGTACTGCACACCTGTGCTACCTAGTGCCTCTCTGGGTCTACTCAGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CACAGCAGCTGATGTGGAAGCTCTCCGGGGAGCAGCGCTGACTGCTGAGCGCGCAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCACTGCGCTCAGACTTCAAGACCATCTTGAC 3451

SULT 12

-09-128-354-1

Sequence 1, Application US/09128354

Patent No. 6337200

GENERAL INFORMATION:

APPLICANT: Morin, Gregg B.

APPLICANT: Geron Corporation

TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants

FILE REFERENCE: 015389-003310US  
CURRENT APPLICATION NUMBER: US/09/128,354  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 08/851,843  
EARLIER FILING DATE: 1997-05-06  
EARLIER APPLICATION NUMBER: US 08/854,050  
EARLIER FILING DATE: 1997-05-09  
EARLIER APPLICATION NUMBER: US 08/911,312  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/912,951  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: US 08/915,503  
EARLIER FILING DATE: 1997-08-14  
EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
EARLIER FILING DATE: 1997-10-01  
EARLIER APPLICATION NUMBER: US 08/974,549  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 08/974,584  
EARLIER FILING DATE: 1997-11-19  
EARLIER APPLICATION NUMBER: US 09/052,864  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 4015  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (36)..(3454)  
OTHER INFORMATION: human telomerase reverse transcriptase (hTERT) cDNA  
US-09-128-354-1

Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260  
Query Match: 59.40% Indels: 52  
DB: Gaps: 13

US-09-042-460-2 (1-1122) x US-09-128-354-1 (1-4015)

Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
Db 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
Qy 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
Db 116 GTGCTGCGCTGGCCACATTTGTGCGGCGCTTGGGCGCCCGAGGCTGGCGGTGGTGAG 175  
Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60  
Db 176 CGCGGGGACCCGCGGCTTCCGCGCGCTGCGCGCTGCGCTGCTGCTGCTGCTGCTGCTG 235  
Qy 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
Db 236 GACGACG 295  
Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
Db 296 GTGGCGCGAGTGTCTGACAGAGGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
Qy 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
Db 356 TTCGCGCTGCTGACG 415  
Qy 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140  
Db 416 ACTTACTTGGCCCAACACCGGTGACCGACCATCTGGGGGAGCGGGGCTGGGGGCTGCTG 475

141 LeuSerArgValGlyAspLeuValTyrLeuLeuValHisCysAlaLeuTyrLeu 160  
146 CTGGCCGCGTGGCGCAGCGTCTGTTTCACTGCTGGCAGCGTGGCGCTTCTTGTG 535  
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnLeuCysAla 180  
536 CTGTGGCTCCAGCTGCGCTTACAGGTGTGGCGCGCGCTGTACAGCTCGCGCT 595  
181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
596 GCCACTCAGGCCCGGCCCGCCACACAGCTAGT---GGACCCGGAAGCGTCTGGGA--- 649  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnLeuLysSerSerArgGlnGlnAlaPro 220  
650 -----TGGACGGGCGCTGGAACTATAGCTCAGGAGCGCGGG 688  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
689 GTCCCGCTGGGCTGCCAGCCCGGCTGGAGAGCGCGGGGAGTGCACCGAAGT 748  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGlnGlyPro--- 259  
749 CTGCCGTTGCCAAGAGCCCGCGCGTGGCGCTCCCTCAGCGGAGCGGCGCGCTT 808  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
809 GGGCAGGGCTCTGGGCGCCACCCGGCGAGGAGCGGTGGACCGAGTACCGTGTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGlnLysAspLeuSerLys 292  
869 GTGGGTTCACCTGCCAGA-----CCCGCCGAAGAGCCACTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GTGGCGCTCTGGCAGCGCGCACTCCACCCATCTGTTGGCGCGCACACCGCGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCCATCCACATCCGCGCCACCGTCTCCGGACACGCTGCTCCCGGGTGTAGCC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCACTTCTTCTCTCTCAGCGC---AAGGAGCAGCTCGCGCTCTTCT 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGlnLillePhe 370  
1091 CTACTCAGCTCTGAGGCCAGCCGCTGCTGGCGCTCGAGGCTCGTGGAGCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTCAGGCCCTGGATGCCAGGACTCCCGCAGGTGCCCCGCTGCCCGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATCGGCCCTGTTCTGGAGCTGCTTGGAACTCCAGCAGCGAGTGGCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGGTCTCTCAAGACGCACTGCCCGCTGGAGTGG-----GTACCCCGCAGCA 1321  
431 Leu-----Asn 432  
1322 GCGGGTGTGTGCGCGGAGAGCCCGCAGGCTCTGTGGCGGCCCGCGAGGAGGAG 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGGCGTGTGAGTGTCTGGCAGCAGCAGCAGCCCGCTGGAGGTATC 1441  
453 GlyPheLeuArgAlaCysLeuValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGCGGGGCTGCTCTGGCGGCTGTGGCCCCCAGGCTCTGGGGCTCCAGGCAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492

1502 AACGACGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGAGCATGCGCAAG 1561  
493 LeuSerLeuGlnLeuLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTGCGAGAGCTGACGTGGAGATGAGCGTGGCGGATGCGCTTGGCTGGCAGG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgLilleAla 532  
1622 AGCCAGGGTGGCTGTGTTCCGGCGCAGAGCACCGTCTGCGTGGAGGATCTTGACC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1682 AAGTTCCTGCTGCTGATGAGTGTGTACGTGCTGCTGAGCTGCTCAGGTCTTTCTTTAT 1741  
553 IleThrGlnSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrpSer 572  
1742 GTACGAGAGACAGCTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgLilleLeuSer 592  
1802 AAGTTGCAAGCATTGGAACTCAGACAGCACTTGAAGAGGCTCAGCTGCGGAGCTGTCG 1861  
593 GlnGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGTCCAGGAGCATCGGGAAGCCCGCTCTGCTGACGTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCTATCCCAAGCTGACGGCTGCGCGCATTTGTGAACATGACTAGCTGCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 552  
1982 AGAACGTTCCGACAGAAAGAGCGCGGCTCTCACCTCAGGGGTGAAGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCTGTCTCACTACGAGCGCGCGCGCGCGCTCTCTGGCGGCTCTGTGTGTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCTGCGCGCATCTGCTGCTGCTGCGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTCAGCTGACTTTGTCAAGGTGGATGTACGCGCGCGCTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTATCGCGCATCATCAAA---CCCCAGAAACAGTACTGCTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyValHisLysSerPheArgArg 752  
2279 CGTGGTATCGCGGTGTCAGAGCGCGCGCATGGCGCAGCTCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGCTCTACTTTCAGACCTCCAGCGGTATCGACAGTTCGTGCTCAGCTGCTGAC 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGCTGCTCATCGACGAGAGCTCCCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGCAGTGGCTCTTCGACGCTTCTTACGCTTCTATGTGCCACCGCGCGTGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGCGCAAGTCTCAGTCCAGTGCAGGAGTCCCGCAGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852





749 CTGCGTTGCCCAAGAGGCGCCAGCGGTGGCGCTCCCTGAGCGCGGAGCGAGCGCCCGTT 808  
260 -----HisArgGlnValLeuProThrProSerGlySerThr--- 272  
809 GGGCAGGGGTCTCGGCCCAACCGCGGAGGAGCGGTGGACCGAGTACCGTGGTTCTGT 868  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
869 GTGGGTGTCACCTGGCAGA-----CCCGCCGAAGAGCCACCTCTTTGGAG 913  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
914 GGTGCGCTCTGGCAGCGCGCACTCCACCCATCCGTGGCGCGCCAGCACACACGGGGC 973  
312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
974 CCCCATCCACATCGCGGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTACGGC 1033  
331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuLeuAsnProSerPhe 350  
1034 GAGCAAGACACTCTCTACTCTCTCAGCGAC---AAGGAGCAGCTGGCGCCCTCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGCCCGCAGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
1151 CTGGGTTTCAGGCGCTTGATCGCAGGAGACTCCCGCAGGTGGTCCCGCTGCCCGAGCG 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTCTGTTCTGAGCTGCTGGGAACACACGCGCAGTCCCGCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
1271 GGGGTGCTCTCAAGACGACTGCGCGCTGCGAGCTCG-----GTACCCCGACGA 1321  
431 Leu-----Asn 432  
1322 GCGGTGTCTGTGCCCGGAGAGCCCGCAGGCTCTGTGGCGGCCCGCAGGAGGAGC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCTGTGTCAGCTGCTCCGCGCAGCACAGCAGCCCTCGCAGGTGAC 1441  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GGTCTGTGCGGCTGCTGCGCGCGCTGGTGGCCCCCAGGCTCTGGGCTCCAGGAC 1501  
473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492  
1502 AACGAACCGCGTCTCTCAGGAACACCAAGAAGTTTCATCTCCCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCTCGCTCAGAGACTGACGTGGAGATGAGCGTGGCGGACTCGGCTTGGCTGGCGAGG 1621  
513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgGlyLeuAla 532  
1622 AGCCAGGGGTGGGTGTCTCCGCGCCGACAGCACCGCTCTGCGTGGAGGATCTCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuLeuArgSerPhePheTyr 552  
1682 AAGTTCCTGACTGGCTGATGAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTCACGAGACACAGCTTTCAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCATTTGAATCAGACGACATTTGAAGAGGGGTGCAGCTGCGGGAGCTGTGCG 1861

593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTTCAGGAGCATCGGAAGCAGCCCGCCCTGCTGACCTCCAGATCCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTTATCCCCAAGCTCAGCGGCTCGGCCGATTTGTGAACATGACTACGTCTGTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGCGCGCAGCTCTCACCTCGAGGGTGAAGGCACCTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGGTGCTCACTACAGCGGGCGCGCGCGCCCGCCCTCTGGCGCTCTGTGTGTGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCGCTGGCGCACTTCTGCTGCTGCTGCGGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGGTGATGTACGGCGCGTACGACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGAACACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCCCTGTGTCAGAAAGCGCCCATGGCGCAGCTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACCTTGACAGACCTCCAGCGCTACATCGCAGTGTGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCGCTGTCATCGACGAGAGCTCCCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGAGTGGCTCTTCGACGCTTCTTACGCTTCTATGTGCCACACCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlnSerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTTACGTCCTCAGTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTTCGACGCTGTCTACGCGGACATGGAGAACAGCTGTTTGGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2692  
873 AlalysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGAAGACAGTGGTGAATCTTCTCTGTGATGATGATGATGATGATGATGATGATGAT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACCGCTTATCCCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGCTGACAGCGACTACTCTCAGCTATGCCCGGACCTTCCATCAGAGCCAGTCTCACC 2932





```

RESULT 15
US-08-912-951-1
; Sequence 1, Application US/08912951
; Patent No. 6475789
; GENERAL INFORMATION:
; APPLICANT: Cech, Thomas R.
; APPLICANT: Lingner, Joachim
; APPLICANT: Nakamura, Toru
; APPLICANT: Chapman, Karen B.
; APPLICANT: Morin, Gregg B.
; APPLICANT: Harley, Calvin
; APPLICANT: Andrews, William H.
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
; THERAPEUTIC METHODS
; NUMBER OF SEQUENCES: 335
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/912,951
; FILING DATE: 14-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002600US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4015 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..3454 /product= "hTERT"
; OTHER INFORMATION: /note= "human telomerase reverse
; OTHER INFORMATION: transcriptase (hTERT) catalytic protein
; OTHER INFORMATION: component"

```

S-08-912-951-1

## Ligment Scores:

Red. No.: 0 Length: 4015  
 core: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservativeness: 122  
 est Local Similarity: 62.36% Mismatches: 260  
 Indels: 52  
 Gaps: 13

S-09-042-460-2 (1-1122) x US-08-912-951-1 (1-4015)

Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuArgSerArgTyrArgGlu 20  
 b 56 ArgCGCGCGCTCCCGCTGCGAGCGTGGCTGCTGCTGCGACCACTACCGCGAG 115  
 Y 21 ValTTPProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40  
 b 116 GTGCTGCGCTGGCCACGCTTGTGCGCGCTGCGCGCCCGACGAGCTGGCGTGGCGAG 175  
 Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTTP 60  
 b 176 CCGCGGAGACCGCGCGCTTCCGCGCTGTGTGCGCGCTGTGTGCGCGCTGTGTGCGCGCTGTG 235  
 Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 b 236 GACGACGCGCGCGCGCGCGCTTCTTCCGCGAGGTGTCTGCTGCTGCTGCTGCTGCTGCTG 295  
 Y 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 b 296 GTGGCGCGAGTGTCTCAGAGGTGTGCGAGCGCGCGCGAGAACGTGTGCTGTGCTGTGCGC 355  
 Y 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerSerValArg 120  
 b 356 TTGCGCTGTGAGCGGCG 415  
 Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTTPMetLeuLeu 140  
 b 416 AGCTACCTGCTCCCAACACAGCTGACCGACCACTGCGGGGAGCGGGGCGTGGCGTGTGCTG 475  
 Y 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 b 476 GTGCGCGCTGTGAGCGGCG 535  
 Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 b 536 CTGTGTGCTCCCGAGCTGCGCTTACAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 595  
 Y 181 ThrThrAspIleTTPProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 b 596 GCGACTCAGCGG 649  
 Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
 b 650 -----TGCGACGGGCTGGAACCATAGCGTCAGGAGGCGCGG 688  
 Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 b 689 GTCCCGCTGGCGCTGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 748  
 Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProAlaGluGluGlyPro--- 259  
 b 749 CTGCGCTTCCCGAGGCGG 808  
 Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTTP--- 272  
 b 809 GGGCAGGGTCTCTGGCGG 868  
 Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 b 869 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCG 913  
 Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysPheProSer 311

Db 914 GGTGCGCTCTCTGCGACGGCCACTCCACCCATCGTGGCGCGCAGCACACCGGGGC 973  
 Qy 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 974 CCCCACATCCACATCGCGGCGCACCGCTCCCTGGGACACAGCTTGTCCCGCGGTGTACGCC 1033  
 Qy 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 1034 GAGACCAAGCACCTCTCTACTCTCTCAGCGAC---AAGAGCAGCTGCGCGCTCTCTTC 1090  
 Qy 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 Db 1091 CTACTCAGCTCTGAGGCCACGCTGCTGCGCTCGGAGCTCGTGAGACCATCTTT 1150  
 Qy 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 Db 1151 CTGGTTCAGGCGCTGGATGCCAGGACTCCCGCAGGTTCGCCCGCTGCCCGCGCGC 1210  
 Qy 391 TyrTTPGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1211 TACTGGCAATGGCGCTCTTCTGGAGCTGTGGGACACACCGCGAGTGCCTTAC 1270  
 Qy 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1271 GGGGTGTCTCAGACGCACTGCCGCTCGAGCTGCG-----GTACCCCGAGCA 1321  
 Qy 431 Leu-----Asn 432  
 Db 1322 GCCGTGTCTGTCGGGAGAGCCCGAGGCTCTGTGCGCGCGCGCGCGAGGAGGAC 1381  
 Qy 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTTPGlnValTyr 452  
 Db 1382 ACAGACCCCGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441  
 Qy 453 GlyPheLeuArgAlaCysLeuCysLysValSerValSerAlaSerLeuTTPGlyThrArgHis 472  
 Db 1442 GGTTCGTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501  
 Qy 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492  
 Db 1502 AACGAACGCGCTTCTCAGGAACACCAAGAGTTCATCTCTCGCGGGAAGCATGCCAAG 1561  
 Qy 493 LeuSerLeuGlnGluLeuMetTTPLysMetLysValGluAspCysHisTTPLeuArgSer 512  
 Db 1562 CTCTCGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1621  
 Qy 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
 Db 1622 AGCCAGGGGTGTGCTGTGTTCCGCGCGCAGACACCGTCTGCTGAGGAGATCTCTGGCC 1681  
 Qy 533 ThrPheLeuPheTTPLeuMetAspThrTyrValGlnLeuLeuArgSerPhePheTyr 552  
 Db 1682 AAGTTCCTGCACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741  
 Qy 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTTPSer 572  
 Db 1742 GTCAGGAGACCATGTTTCAAGAACAGAGCTCTTTTCTACCGGAGAGTGTCTGAGC 1801  
 Qy 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
 Db 1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGCGAGCTGCGGAGCTGTG 1861  
 Qy 593 GlnGluGluValArgHisGlnAspThrTTPLeuAlaMetProIleCysArgLeuArg 612  
 Db 1862 GAAGCAGAGGTGAGCGAGCTGCGGAGAGCGCGCGCGCTGTGCTGCTGCTGCTGCTGCTGCTG 1921  
 Qy 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
 Db 1922 TTCATCCCAAGCTGAGCGGCTGCGCGCGAGTGTGAACATGAGCTACGTGCTGGAGGC 1981  
 Qy 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652

1982 AGAAGCTTCGACAGAAAAGAGGCGGAGCGCTCTCACCTCGAGGCTGACGACACTGTC 2041  
553 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCAACTACAGAGCGGCGCGCCCGCCGCTCTCTGGCGCTCTGTGTGGCC 2101  
673 MetAsnAspIleTyrArgThrTTPArgAlaPheValLeuArgValAlaLeuAspGln 692  
2102 CTGACGATATCCACAGGCGCTGGCGACCTCTGTGTCTGGTGTGGGCGCCAGGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGACTTTCACAGGTGGATGTACGGGCGGTACGACACCATCCCCCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTCACGAGGTCTATCGCCAGCATCATCAA- -CCCAGAACACGTACTGCGTG 2278  
733 ArgGlnTyrAlaValValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATCCGTGTGTCACAAAGGCGGCCCATGGGCACGTCCGCAAGGCCCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACCTTGACAGACTCCAGCGGTATACGACAGTTCGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC- -AGCCCGCTGAGGGATGCGCTCGTTCATCGACGACAGGTCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
2453 GAGCCAGAGTGGCTCTTCGACGTCTTCCTACGCTTCATGTGCCACACCGCCGTGGCC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTACGTCCAGTCCAGGGGATCCCGACGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGCAGCTGTGCTACGGCAGCATGGAGAACAAAGCTTTTGGCGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTGCTCTCGCTTGTGGTGGATGATTTCTGTGGTACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAAACCTTCCTCAGGACCTGGTCCGAGGTGTCCCTGAGTATGGCTGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyValAlaAlaPro 912  
2753 TTGCGGAAGACAGTGGTGAACCTTCCTGTAGAGACGAGGCGCTGGGTGGACGGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGGCGCCACGCGCTATTCCCTGGTGGCGCTGCTGTGGATACCCGAC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGACGAGCGACTACTCCAGTATGCGCGGACCTCCATCAGACCGAGTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGAGGACATCGTCCGAAACTCTTTGGGGTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACACGCTGTTTCTGGATTTGCAAGGTGAACAGCGCTCCAGACGGGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAGATCTCTCTGCTGCGAGGCGTACAGGTTTTCACGCAATGTGTGTCGACCTC 3112

1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
3113 CATTTCATCAGCAAGTTTGAAGNACCCACATTTTCTGCGCGTCATCTCTGACACG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCCTCCCTCTGTACTCCATCTGAAGCCAGAACGACGAGGATGTGCTGGGGGCCAAG 3232  
1053 GlySer- - - - -PheProProGluAlaAlaHisIleLeuCysTyrGlnAlaPheLeu 1069  
3233 GGCGCGCGCGCCCTCTGCCCTCCGAGGCCGTGAGTGGCTGTGCCACCAAGCATTTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGTGTACCTACGTGCCACTCTCTGGGGTCACTCAGGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
3353 CAGCGCAGCTGAGTCGAAGCTCCCGGGGACGAGTGTGCTGCCCTGGAGGCCGACGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCGGCACTGCCCTCAGACTTCAAGACCATCCTGGAC 3451

Search completed: February 1, 2004, 19:35:25  
Job time : 360 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

n on: February 1, 2004, 18:05:20 ; Search time 877 Seconds  
(without alignments)  
4662.705 Million cell updates/sec

tle:  
xrefact score: 5901  
quence: 1 MTRAPRCPAVRSILRSRYRE.....TILKAADPALSTDFQITLD 1122

oring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Xgapop 10.0, Xgapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

arched:  
2434939 seqs, 1822278265 residues

tal number of hits satisfying chosen parameters: 4869878

nimum DB seq length: 0  
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

mmmand line parameters:  
MODEL=frame+p2n.model -DEV=xlh  
=/cgn2\_1/USPTO.spool/US09042460/runat\_01022004\_150553\_17127/app\_query.fasta\_1.1287  
B=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
COPI=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
RANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
HR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0  
TAXLEN=200000000 -USER=US09042460@cgn 1.1 586 @runat\_01022004\_150553\_17127  
ICPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
JONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

atabase : Published Applications NA:\*

1:	/cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17:	/cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Query	Score	Match	Length	ID	Description
-------	-------	-------	--------	----	-------------

1	3505	59.4	3396	10	US-09-749-728B-32	Sequence 32, Appl
2	3505	59.4	4015	9	US-09-733-294A-3	Sequence 3, Appl
3	3505	59.4	4015	10	US-09-990-080-1	Sequence 1, Appl
4	3505	59.4	4015	10	US-09-843-676-224	Sequence 224, App
5	3505	59.4	4015	10	US-09-953-052-1	Sequence 1, Appl
6	3505	59.4	4015	15	US-10-053-758-224	Sequence 224, App
7	3505	59.4	4015	15	US-10-208-243-1	Sequence 1, Appl
8	3505	59.4	4015	15	US-10-054-295-224	Sequence 224, App
9	3505	59.4	4015	15	US-10-054-611-224	Sequence 224, App
10	3505	59.4	4015	15	US-10-105-963-1	Sequence 1, Appl
11	3505	59.4	4015	15	US-10-044-692-1	Sequence 1, Appl
12	3505	59.4	4015	15	US-10-044-692-1	Sequence 1, Appl
13	3498	59.3	4027	12	US-10-385-882-1	Sequence 1, Appl
14	3496	59.2	13766	13	US-10-105-616-1	Sequence 1, Appl
15	3496	59.2	13766	13	US-10-105-616-1	Sequence 6, Appl
16	3473	58.9	8742	13	US-10-105-616-6	Sequence 173, App
17	3287	55.7	4029	11	US-09-843-676-173	Sequence 173, App
18	3287	55.7	4029	11	US-09-438-486-173	Sequence 173, App
19	3287	55.7	4029	15	US-10-053-758-173	Sequence 173, App
20	3287	55.7	4029	15	US-10-054-295-173	Sequence 173, App
21	3287	55.7	4029	15	US-10-054-611-173	Sequence 173, App
22	3276.5	55.5	3855	15	US-10-044-692-4	Sequence 4, Appl
23	3276.5	55.5	3855	15	US-10-044-692-4	Sequence 4, Appl
24	2132	36.1	1866	15	US-10-294-778-11	Sequence 11, Appl
25	1844.5	31.3	2176	15	US-10-044-692-3	Sequence 3, Appl
26	1844.5	31.3	2176	15	US-10-044-692-3	Sequence 3, Appl
27	1786.5	30.3	2171	10	US-09-843-676-100	Sequence 100, App
28	1786.5	30.3	2171	10	US-09-766-253-100	Sequence 100, App
29	1786.5	30.3	2171	11	US-09-438-486-100	Sequence 100, App
30	1786.5	30.3	2171	15	US-10-053-758-100	Sequence 100, App
31	1786.5	30.3	2171	15	US-10-054-295-100	Sequence 100, App
32	1786.5	30.3	2171	15	US-10-054-611-100	Sequence 100, App
33	1511.5	25.6	1314	15	US-10-294-778-9	Sequence 9, Appl
34	1497	25.4	1311	15	US-10-294-778-1	Sequence 1, Appl
35	1336	22.6	4200	15	US-10-044-692-6	Sequence 6, Appl
36	1336	22.6	4200	15	US-10-044-692-6	Sequence 6, Appl
37	1336	22.6	15418	9	US-09-783-203-1	Sequence 1, Appl
38	1336	22.6	15418	10	US-09-934-427A-1	Sequence 1, Appl
39	1336	22.6	15418	11	US-09-935-419A-1	Sequence 1, Appl
40	1336	22.6	15418	15	US-10-141-220-1	Sequence 1, Appl
41	1336	22.6	15418	15	US-10-023-969-1	Sequence 1, Appl
42	1336	22.6	15418	15	US-10-206-447-1	Sequence 1, Appl
43	1336	22.6	51552	9	US-09-733-294A-30	Sequence 30, Appl
44	972.5	16.5	873	13	US-10-282-960-1	Sequence 1, Appl
45	702	11.9	519	13	US-10-282-960-2	Sequence 2, Appl

## ALIGNMENTS

## RESULT 1

US-09-749-728B-32

; Sequence 32, Application US/09749728B

; Patent No. US20020142457A1

; GENERAL INFORMATION:

; APPLICANT: Umezawa, Akihiro

; APPLICANT: Hata, Jun-Ichi

; APPLICANT: Fukuda, Keiichi

; APPLICANT: Ogawa, Satoshi

; APPLICANT: Sakurada, Kazuhiro

; APPLICANT: Gojo, Satoshi

; APPLICANT: Yamada, Yoji

; TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIO

; FILE REFERENCE: 00766.000043

; CURRENT APPLICATION NUMBER: US/09/749,728B

; CURRENT FILING DATE: 2001-09-17

; PRIOR APPLICATION NUMBER: H11-372826

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: PCT-JP00-01148

; PRIOR FILING DATE: 2000-02-28

; PRIOR APPLICATION NUMBER: PCT-JP00-07741

; PRIOR FILING DATE: 2000-11-02

; NUMBER OF SEQ ID NOS: 80

; SOFTWARE: Patent in Ver.2.0

```

SEQ ID NO 32
LENGTH: 3396
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
OTHER INFORMATION: (1)..(3399)
3-09-749-728B-32

Ligment Scores:
red. No.: 0
score: 3505.00
percent Similarity: 72.94%
set Local Similarity: 62.36%
very Match: 59.40%
B: 10
Gaps: 13

S-09-042-460-2 (1-1122) x US-09-749-728B-32 (1-3396)

Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
b 1 ATGCGCGCGCTCCCGCTGCCGAGCGCTGCGCTGCCGCTGCCGCTGCCGCTGCCGCTGCCG 60
Y 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
b 61 GTGCTGCGCTGGCCACGCTTCTGCGCGCGCTGCCGCGCGCTGCCGCGCGCTGCCGCTGCCG 120
Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
b 121 CGCGGGGACCGCGCGCTTTCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTG 180
Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
b 181 GAGCGACGCGCGCGCGCGCGCTTCTGCGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCG 240
Y 81 ValAlaArgValGlnArgLeuCysGluArgGlnGlnArgGlnGlnArgGlnGlnArgGln 100
b 241 GTGGCGCGAGTGTGCAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Y 101 PheGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
b 301 TTCGCGCTGCTGGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
b 361 AGCTACCTGCCCAACACAGCTGACCGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Y 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
b 421 CTGCGCGCGTGGCGGACGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
b 481 CTGTGTGGCTCCCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 540
Y 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
b 541 GCACCTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
b 595 -----TGCAGCGCGCTTGAACCATAGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCG 633
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
b 634 GTCCCGCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 693
Y 241 ValProSerAlaLysLeuAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
b 694 CTGCGCTGCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 753
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272

```

```
613 PheileProLysProAenglyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1867 TTCATCCCAAGCTGACGGCTGCGGCCGATTTGTGAACATACGCTACGCTGGGAGCC 1926
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1927 AGAAGCTTCCGACAGAAAGAGAGGCGGAGCGTCTACCTCGAGGTGAGGACACTGTC 1986
653 SerMetLeuAenTyrGluArgThrLysHisProHisLeuMetGlySerValLeuGly 672
1987 AGCGTGCTCAACTACGAGCGGCGCGCGCGCCCTCTCTGGCGCCTCTGTGCTGGGC 2046
673 MetAenAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2047 CTGACGATATCCACAGGCTGCGGCACCTTGTGTGCTGTGTGCGGCCAGGACCG 2106
693 ThrProArgMetTyrPheValAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2107 CCGCTGAGCTGTACTTGTCAAGGTGATGTGACGGGCGGTACGACACCATCCGCCAG 2166
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2167 GACAGGCTCAGGAGGTATCGCCAGCATCATCAAA---CCCGAACACGCTACTCGGTG 2223
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2224 CGTGGTATGCGGTGTCAGAGCGGCCCATGGGCACGTCCGCAAGCGCTTCAAGAGC 2283
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2284 CACGTCTCTACTCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTCACCTGCAG 2343
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792
2344 GAGACC-----AGCCCGCTGAGGAGTGGCGTGTATCGACGAGAGCTCCCTCGAT 2397
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812
2398 GAGGCGACAGTGGCTCTTCAGAGCTCTTCATGCTTCACTGCGCACCAAGCGCGTGGC 2457
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2458 ATCAGGCGCAAGTCTCTAGCTGCGTGGCAGGAGATCCCGAGGCTCCATCTCTCCAG 2517
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2518 CTGCTCTCAGCGCTGTGTACGGGCGACATGAGAGCAAGCTGTTTGGCGGATTCGGCG 2577
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872
2578 GACGGGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2637
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2638 GCGAAACCTTCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2697
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyAlaAlaPro 912
2698 TTGCGGAAGACAGTGGTGAATCTCCCTGTAGAAGACAGAGCCCTGCTGCTGCTGCTGCT 2757
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2758 GTTCAGATGCGGCGCCAGCGCTATTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2817
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2818 CTGAGGTGCGAGAGCGACTACTCCAGCTATGCCCGGACCTCCATCAGAGCGAGCTCCAC 2877
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
2878 TTCACCGCGGCTTCAAGGCTGGAGGAAACATGCTGCGCAACTCTTTTGGGCTTGGCG 2937
```

```
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
Db 2938 CTGAGGTGCACAGCTGTTCGTGATTTGCAAGTGAACAGCTCCAGCGTGTGCACC 2997
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
Db 2998 AACATCTACAAGATCCTCTGCTGTCAGGGGTACAGTTTCAAGCATGTGTGTGCAGCTC 3057
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3058 CCATTTTCATCAGCAAGTTTGGAGAACCCCATTTTCTCGCGCTCATCTCTGACAGC 3117
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
Db 3118 GCCTCCTCTGCTACTCCATCTGTAAGCCCAAGACGAGGATGTCTGGGGGCCAAG 3177
QY 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069
Db 3178 GCGCGCGCGCGCTCTGCTGCTGCGAGCGGTGCGAGTGGCTGTGCCCAAGCATTTCTCTG 3237
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
Db 3238 CTACAGCTGACTCGACACCGTGTACCTACCTAGTCCACTCTGGGGTCACTCAGGACGCC 3297
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
Db 3298 CAGACGCGAGTGAAGTCCCGGGGAGCGAGCTGACTGCTGAGGCGCGAGGCC 3357
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3358 AACCGCGACTGCTCCCTCAGACTTCAGAGCACTCTGGAC 3396

RESULT 2
US-09-733-294A-3
; Sequence 3, Application US/09733294A
; Patent No. US20020045588A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: William Gaarde
; APPLICANT: Susan M. Freier
; APPLICANT: Edward V. Wanciewicz
; TITLE OF INVENTION: ANTISENSE MODULATION OF TERT EXPRESSION
; FILE REFERENCE: ISPH-0527
; CURRENT APPLICATION NUMBER: US/09/733,294A
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/572,423
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 108
; SEQ ID NO 3
; LENGTH: 4015
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)...(3454)
US-09-733-294A-3

Alignment Scores:
Pred. No.: 0 Length: 4015
Score: 3505.00 Matches: 719
Percent Similarity: 72.94% Conservative: 122
Best Local Similarity: 62.36% Mismatches: 260
Query Match: 59.40% Indels: 52
DB: 13 Gaps: 13

US-09-042-460-2 (1-1122) x US-09-733-294A-3 (1-4015)
QY 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 56 ATCCGCGCGCTCCCGCTGCGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 115
QY 21 ValTrpProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgLeuValGln 40
```

b 116 GTGTCGCGCTGGCCACAGTTCGTGCGCGCGCTGGGGCCCCCAGGGCTGGCGCTGGTCAG 175  
y 41 ProGlyAspProLysIleThrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
b 176 CGCGGGACCGCGCGCTTCGCGCGCTGCTGGCCAGTGCCTGTGTGCGTGGCCCTGG 235  
y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80  
b 236 GACGACGCGCGCGCGCGCGCTTCCTTCGCGCAGGGTCTTCCTCCCAAGAGAGCTG 235  
y 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
b 296 GTGGCCCGAGTCTGCAGAGGCTGTGCGAGCGCGCGGAGAACGTCGTGGCCTTCGGC 355  
y 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTGCGGCTCTGCAGCGCGCGCGCGCGCGCTTCCTTCGCGCAGGGTCTTCCTCCCAAGAGAGCTG 415  
y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACTGCCCCACACGGTGACCGACGACTGCGGGGAGCGGGGCTGGCTGCTG 475  
y 141 LeuSerArgValIleAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 CTGCGCGCGCTGGCGACACGTCGTGTTCACCTGCTGGCACGCTGCGGCTCTTTGTG 535  
y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGCTGCTCCACGTCGCCCTACGAGTGTGCGGCGCGCGCTGTACCGACTCGGCGCT 595  
y 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCCACTCAGCGCGCGCGCGCGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 649  
y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TGCGAGCGGGCTCGGACCTAGCGTCAGGGAGCGCGCG 688  
y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerSer 240  
b 689 GTCCCCCTGGGCTGCCAGCCCGCGGTGGAGAGCGCGCGCGCGCGCGCGCGAGT 748  
y 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTGCGCTGCTCCCAAGAGCGCGCGCGCTGCGCTGCGCTGCGCGCGCGCGCGCGCTT 808  
y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
b 809 GGCAGGGGTCTGGGCCCCACCGCGCGAGCGCTGGACCGAGTGCCTGTTCTGT 868  
y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTCACTGCCAGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 913  
y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GTGCGCTCTCTGGCAGCGCGCGCTCCACCCATCCGTTGGCGCGCGCGCGCGCGCG 973  
y 312 SerThrSerLeuLeuSerProProArgGlnInAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCGCATCCATCGCGCGCGCGCGCTGCTGGCACGCGCTTGTCCCGCGCTGTAGCC 1033  
y 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACCTTCCTCTACTCTCTCAGCGGAC---AAGAGCAGCTCGGCGCTCTTC 1090  
y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTCTGAGCGCGCGCGCTGCTGCGCTCGGAGCTCGTGAGACCATCTTT 1150  
y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
b 1151 CTGGGTTCCAGCGCGCGCTGATGCCAGGAGTTCGCCCGCGAGTTCGCCCGCTGCCCGAGCGC 1210

QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
Db 1211 TACTGGCAATATGGGCGCGCTTCTTCTGGAGCTCTTGGGAACACCGCGAGTGGCCCTAC 1270  
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
Db 1271 GGGTGTCTCTCAAGACGCACTGCCGCTGCGAGTGGC-----GTCACCCCGAGCA 1321  
QY 431 Leu-----Asn 432  
Db 1322 CCCGCTGTCTGTGCCCGGAGAACCCCGAGGCTCTGTGGCGCGCCCCCGAGGAGGAGAC 1381  
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
Db 1382 ACAGACCCCGCTGCTGTGTCAGCTGCTCCGCCAGCAGCAGCGCCCTGGCAGGTGTAC 1441  
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GGTTCGTGCGGGCTGCTGCGCGCGCTGTCGCCCGCGCGCTCTGGGGCTCCAGGCAC 1501  
QY 473 AsnGluArgArgPhePheLysAsnLeuLysPheLysLeuGlyLysTyrGlyLys 492  
Db 1502 AACGACCGCGCTTCTCAGAACACCGAGAGTTTATCTCCCTGGGAGAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTGCAGGAGCTGACGTGGAAGTGGCGTGGAGCTGCGCTTGGCTGCGCAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCGAGGGTGGCTGTGTTCGCCCGCAGACACCGCTCTGCGTGAGAGATCTCTGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1682 AAGTTCCTGCACTGGCTGATGAGTGTCTAGCTGCTGAGCTGCTCAGGCTCTTCTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTACCGGAGACACGCTTTCAAAGAACAGAGCTCTTTTCTACCGGAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AGTTGCAAGCANTGGANATCAGACACACTTGAAGAGGGTGCAGCTGCGGAGAGTGTG 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTGAGGACGATCGGGAAGCCAGCGCGCGCTGCTGAGCTCCAGACTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCATCCCAAGCTTGACGGCTCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAACGTTCCGACGAGAAAGAGCGCGCGCTCTACCTCGAGGGTGAAGGACCTGTTTC 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCTGCGCGCTCTGCTGCTGCTGCTG 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGAGCATATCCACAGGGCGCTGGCGCACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGTACTTGTCAAGGTGTGTGCGGCGCGCTGACGACCATCCCGCCAG 2221  
QY 713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGGCTCAGCGGAGTTCATCGCCAGCATCATCAAA---CCCCAGAACAGTACTGGTG 2278



161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGCTGGCTCCAGCTGCGCTACAGGTGCGGGCGCGCTGTACAGCTCGCGCT 595  
y 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCACCTCAGCG 649  
y 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerArgGlnGlnAlaPro 220  
b 650 -----TGCGAACCGCGCTCGGAACCATAGCGTCAGGAGCGCGG 688  
y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
b 689 GTCCCGCTCGCGCTCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748  
y 241 ValProSerAlaLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTGCGCTTGCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTT 808  
y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272  
b 809 GCGCAGGGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTT 868  
y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTACCTGCCAGA-----CCCGCGGAGAGCGCACCTCTTTGGAG 913  
y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GGTGGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGG 973  
y 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCCCATTCCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGTAG 1033  
y 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1090  
y 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1150  
y 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
b 1151 CTGGGTTCCAGCG 1210  
y 391 TyrTrpGlnMetArgProLeupheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
b 1211 TACTGGCAATCGCT 1270  
y 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
b 1271 GGGGTCTCTCTCAAGCAGCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1321  
y 431 Leu-----Asn 432  
b 1322 GCGGTGTCTGTGCGCGCGGAGAGCCCGCGCGCTCTGTGGCGCGCGCGCGCGAG 1381  
y 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
b 1382 ACAGACCGCGCTGCTGTGTGAGTGTCTCGCGCAGCAGCAGCGCGCGCGCGGT 1441  
y 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
b 1442 GGCTTCTGCGCGCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501  
y 473 AsnGluArgArgPhePheLysAsnLeuLysPheIleSerLeuGlyLysTyrGlyLys 492  
b 1502 AACGAACCGCGCTTCTCTCAGGAACACCAAGAGTTCATCTCTCTCTGGGAGCATGCCAAG 1561  
493 LeuSerLeuGlnGlnLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
b 1562 CTCTCGCTCAGAGCTGACGTGGAAGATGACGTGCGGAGCTGCGCTTGGCTGCGCAGG 1621  
y 513 SerProGlyLysAspArgValProAlaGluHisArgLeuArgGluArgIleLeuAla 532  
b 1622 AGCCAGGGGTGCTGCTGCTGCGCGCGCAGAGCAGCGTCTCGTGAGAGATCTCGGCC 1681  
y 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
b 1682 AAGTTCCTGACCTGCGCTGATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
y 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
b 1742 GTACGAGACACAGCTTCAAGAGACAGCTCTTTCTACCGGAAGAGTGTCTGGAGC 1801  
y 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
b 1802 AAGTTGCAAGCATTTGAATCAGACAGCACTTGAAGAGGTGACAGCTGCGGAGCTGCG 1861  
y 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
b 1862 GAACGAGGTGAGCAGCATCGGGAAGCGCGCGCGCTGCTGAGCTCCAGACTCCGC 1921  
y 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
b 1922 TTATCCCCAAGCTGACGGCTGCGCGCGCTGTAACATGAGCTAGCTGCTGGGAGCC 1981  
y 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
b 1982 AGAACGTCGCGAGAGAAAGAGCGCGCGCGCTCTCCTCGAGGGTGAAGCACTGTTC 2041  
y 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
b 2042 AGCTGTCTCAACTACGAGCGCGCGCGCGCGCGCTCTCTGCGCGCTCTGCTGGGC 2101  
y 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
b 2102 CTGACCATATCCAGAGGCGCTGCGCACCTCTGCTGCTGCTGCTGCGCGCGCGAGCCCG 2161  
y 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValArgAspAlaIleProGln 712  
b 2162 CCGCTGAGCTGTACTTGTCAAGGTGATGTGACGGCGCGCTAGCACCATCCCGCAG 2221  
y 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
b 2222 GACAGGCTCAGGAGGTATCGCGCAGCATCATCAA---CCCGAAGAACAGTACTCGGTG 2278  
y 733 ArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
b 2279 CGTGGTATGCGGTGTCTCAGAGCGCGCGCGCGCGCTCCGCAAGCGCTTCAAGAGC 2338  
y 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
b 2339 CACGTCTCTACTTACAGACCTCCAGCGCTATATGCGACAGTCTGCTGCTCACCTCCG 2398  
y 773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
b 2399 GAGAC-----AGCCCGCTGAGGAGTCCGTCTATCGAGCAGAGCTCTCTCTGAT 2452  
y 793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
b 2453 GAGCGCAGCAGTGGCTCTTCGACGTCTCTACGCTTCTATGTGCCACCGCGCTGCGC 2512  
y 813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnLysSerSerLeuSerThr 832  
b 2513 ATCAGGGCAAGTCTTACGTCCAGTCCAGGAGTCCCGCAGGGCTCCATCTCTCCAGC 2572  
y 833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
b 2573 CTGCTCTGAGCTGTCTACGCGGACATGAGACAAGCTGTTCGCGGGATTCGCGCG 2632  
y 853 AspGlyLeuLeuArgPheValAspAspPheLeuValThrProHisLeuAspGln 872





b 296 GTGGCCCGAGTCTGTCAGAGGCTGTGGCGCGCGCGAGAACGTCGTGGCTTCGGC 355  
y PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTCCGGCTGCTGACAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
y SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACCTGCTCCCAACACGGTGACCGACGCTGTGGGGGAGCGGGGTGGGGCTGCTG 475  
y LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 GTGGCGCGCGTGGGGGACACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 535  
y LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGGTGGCTCCACGCTGCGCCCTACCAAGTGTGCGCGCGCGCGCGCGCGCGCGCG 595  
y ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCCACTCAGCG 649  
y AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TGGCAACGGCGCTGGAACCATAGCGTCAGGAGCGCGCG 688  
y LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
b 689 GTCCCTCGGGCTGCGACCGCGCGGTGCGAGAGGCGCGCGCGCGCGCGCGCGCG 748  
y ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
b 749 CTCCCTGTCCTCCCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
y -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
b 809 GCGCAGGGTCTCTGGCGCCACCGCGCGCAGCGCTGGACCGAGTGACCGGTTCGT 868  
y ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
b 869 GTGGTGTCTACCTCCACGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 913  
y GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GTGGCTCTCTGGCAGCGCGCGCTCCACCCATCCGTGGCGCGCGCGCGCGCGCG 973  
y SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
b 974 CCCCACATCCACATCGCGCGCACCGCTCCCTGGGACACCGCTGTCTCCCGCGGTGTACGCC 1033  
y GlnThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
b 1034 GAGACCAAGCACATCTCTACTCTCAGCGAC---AAGGACAGCTGGCGCCCTCTTC 1090  
y LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
b 1091 CTACTCAGCTCTGTAGCGCCACGCTGCTGCGCTCGGAGCTCGTGAGACCATCTTT 1150  
y LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
b 1151 CTGGGTTCACAGCGCGCTGGATCCAGGAGCTCCCGCAGGTTGCCCGCGCTGCCCGCGC 1210  
y TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
b 1211 TACTGGCAATCGCGCCCTGTCTGTGGAGCTGTGGGAACCAACCGCGAGTGCCCTTAC 1270  
y ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
b 1271 GGGGTGCTCTCAAGAGCGCACTGCCCGCTGCGAGCTGCG-----GTCAACCCACGA 1321  
y Leu-----Asn 432  
b 1322 CCGCGTGTCTGTGCCCGGAGAGCCCGCGGCTCTGTGGCGCGCGCGCGCGCGCGAGGAC 1381

QY 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
Db 1382 ACAGACCCCGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1441  
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
Db 1442 GCGTTCGTGGGGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1501  
QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyLysTyrGlyLys 492  
Db 1502 AACGAACCGCGCTTCCTCAGAAACACCAAGAGATTCTTCCTCTGGGGAAGCATGCCAAG 1561  
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
Db 1562 CTCTCGCTCAGGAGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGGCGAGG 1621  
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
Db 1622 AGCCGAGGGGTGGCTGTGTTCGCGCGCGAGACCGCTCTGGTGAGGAGATCTCTGGCC 1681  
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
Db 1682 AAGTTCCTGCATCGCTGATGAGTGTCTGCTCGAGCTGCTCAGGTCTTCTTTAT 1741  
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
Db 1742 GTCAAGGAGACACGCTTTCAAAAGAACAGAGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
Db 1802 AAGTTCGAAGCATTTGAAATCAGACACACTTGAAGAGGGTGCAGCTGCGGAGCTGTG 1861  
QY 593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
Db 1862 GAAGCAGAGGTTCAGGACATCGGGAAGCCAGCGCGCGCTGCTGACGTCCAGACTCCGC 1921  
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
Db 1922 TTCTATCCCAAGCTTCAGCGGCTGCGCGCATGTGTGAACATGAGTACGTCTGGGAGCC 1981  
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
Db 1982 AGAACGTTCCGACAGAAAGAGCGCGCGCGCTCTCACCTCGAGGGTGAAGGACACTGTT 2041  
QY 653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
Db 2042 AGCGTCTCAACTACGAGCGCGCGCGCGCGCGCGCTCTGCGCGCTCTGTGTGGGC 2101  
QY 673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
Db 2102 CTGGACGATATCCACAGCGCTGGCGACCTTCTGCTGCTGCTGCGCGCGCGCGCGCG 2161  
QY 693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
Db 2162 CGCGCTGAGCTGATCTTGTCAAGTGGATGTGAGCGCGCGCGCGCGCGCTCTGTGTGG 2221  
QY 713 GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
Db 2222 GACAGGCTCAGGAGTCTATCGCGACGATCATCAA---CCCCAGAACGCTACTGCGTG 2278  
QY 733 ArgGlnTyrAlaValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
Db 2279 CGTCGGTATGCGGTGTCAGAGCGCGCGCGCGCGCGCGCTGCCCAAGGCTTCAAGAGC 2338  
QY 753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
Db 2339 CAGCTCTTACCTTTCAGACAGCTCCAGCGCGTACATGCGACAGTTCGTGGCTCACCTCAG 2398  
QY 773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
Db 2399 GAGACC-----AGCCCGCTGAGGGATCCGCTGCTCATCGACGAGAGCTCTCTCCCTGAAT 2452

793 G1uSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValVallys 812  
||||:|||||  
2453 GAGCCAGAGAGTGGCTCTTCGACGCTCTCCACGCTTATGTCGCACACGCCCTGGCC 2512  
||||:|||||  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyLeuProGlnGlySerSerLeuSerThr 832  
||||:|||||  
2513 ATCAGGGGCAAGTCTACGTCACGTCAGTCAGGAGATCCCGAGGGTCTCCATCTCTCCACG 2572  
||||:|||||  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
||||:|||||  
2573 CTGCTCTGCGAGCTGTCTACGCGCAGATGGAGAACAGCTGTTGCGGGGATTGCGCGG 2632  
||||:|||||  
853 AspGlyLeuLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
||||:|||||  
2633 GACGGGCTGCTCTCGGTTGGTGGATGATTTCTTGTGTGACACCTCACCTCACCCAC 2692  
||||:|||||  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
||||:|||||  
2693 GCGAAACCTTCTCCAGGACCTCGGTCCGAGGTGTCCCTGATGATGCTGGTGGTGAAC 2752  
||||:|||||  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyValAlaAlaPro 912  
||||:|||||  
2753 TTGCGGAGACAGTGTGATCTCTCTAGAGACGAGCCCTGGTGGCAGCGCTTTT 2812  
||||:|||||  
913 TyrGlnLeuProLalaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
||||:|||||  
2813 GTTCAGATGCGCGCCACGCGCTATTCCCTGCTGGTGGCTGCTGTGTGATACCGGACC 2872  
||||:|||||  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
||||:|||||  
2873 CTGGAGGTGAGAGCGACTACTCCAGTATGCCGAGCTCTCCATCAGACCCAGTCTCACC 2932  
||||:|||||  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
||||:|||||  
2933 TTCAACCGCGCTTCAAGGCTGGAGGAACTGCTGCGAACTCTTTGGGCTCTTGGG 2992  
||||:|||||  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
||||:|||||  
2993 CTGAAGTGTACAGCGCTTTCTTGGATTTGCAAGTGAACAGGCTCCAGACGGTGTGCACC 3052  
||||:|||||  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
||||:|||||  
3053 AACATCTACAGATCTCTCTGCTGACGGGTACAGTTTCAGCATGTGTGCTGCAGCTC 3112  
||||:|||||  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
||||:|||||  
3113 CCATTTTCATCAGCAAGTTTGGAAAGAACCCACATTTTCTCGCGGTCTATCTCTGACACG 3172  
||||:|||||  
1033 AlaserCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
||||:|||||  
3173 GCCTCCCTCTGTACTCTCATCTCTGAAGCCAGACGAGGATGTGCTGGGGCCCAAG 3232  
||||:|||||  
1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
||||:|||||  
3233 GCGCGCCGCGCCCTCTGCTCCCTCCGAGCGGTGACGTGCTGTGCGCCCAAGCATTCCTG 3292  
||||:|||||  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
||||:|||||  
3293 CTCAAGCTGACTGACACCGTGTCACTAGTGCCTCTCTGGGTCTCAGGACAGCC 3352  
||||:|||||  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
||||:|||||  
3353 CAGACGAGCTGAGTCGGAAGTCTCCCGGGAGCAGCGTCACTGCTGCGAGCGCCAGCC 3412  
||||:|||||  
1110 AspProLalaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
||||:|||||  
3413 AACCCGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3451  
||||:|||||

:SULT 5

:-09-953-052-1

Sequence 1, Application US/09953052

Patent No. US20020173476A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin B.  
Andrews, William H.  
TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/953,052  
FILING DATE: 14-Sep-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/052,919  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4015 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..3454  
OTHER INFORMATION: /product= "human telomerase reverse transcriptase (hTERT)"  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-953-052-1

Alignment Scores:  
 Seq. No.: 0 Length: 4015  
 core: 3505.00 Matches: 719  
 Percent Similarity: 72.94% Conservative: 122  
 Best Local Similarity: 62.36% Mismatches: 260  
 Query Match: 59.40% Indels: 52  
 B: 10 Gaps: 13

S-09-042-460-2 (1-1122) x US-09-953-052-1 (1-4015)

```

Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
b 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCACTACCGCGAG 115
Y 21 ValTyrProLeuAlaThrPheValArgArgLeuGlyProGluGlyArgArgLeuValGln 40
b 116 GTGCTGCGCTGCGCAGCTTGTGCGCGCGCTGCGCGCCCGAGGGCTGGCGCTGGTGCAG 175
Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTyr 60
b 176 CGCGGGGACCGCGCGCTTCCGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 235
Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
b 236 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295
Y 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
b 296 GTGGCGCGAGTGTGCAGAGGTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 355
Y 101 PheGluLeuLeuAsnGluAlaArgGlyClyProProMetAlaPheThrSerSerValArg 120
b 356 TTGCGCTGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTyrMetLeuLeu 140
b 416 AGCTACTGCTCCCAACCGGTGACCGACGCTGCGGGGAGCGGGGCTGGGGCTGCTG 475
Y 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
b 476 GTGCGCGCGCTGGCGAGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
b 536 GTGCTGCTCCAGCTGCGCTTACAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 595
Y 181 ThrThrAspIleTyrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
b 596 GCCTACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
b 650 -----TGCAGACGGGCTGTGAACCATACGCTCAGGAGCGCGGG 688
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
b 689 GTCCCTCTGGGCTCCAGCCCGCGGTGCGAGAGCGCGCGGGCGAGTGCAGCGCGAGT 748
Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
b 749 CTGCGCTTCCCAAGAGGCCCGAGCGTGTGCGCTGCGCGCGCGCGCGCGCGCGCGGTT 808
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTyr--- 272
b 809 GGGCAGGGCTCTGGGCCACCGCGCGCGAGCGCTGTGACCGAGTGTGCTGCTGCTGCT 868
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
b 869 GTGGTGTACCTGCCAGA-----CCCGCCGAAGAAGCGCACCTCTTTGGAG 913
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
b 914 GGTGCGCTCTGTGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
  
```

```

QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 974 CCCCATTCCATCGCGGCGCACCATCTCCCTGGGACACGCTTGTCCCGCGGTGTACGCC 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
DB 1034 GAGACCAAGCAGCTTCTCTACTCTCTCAGGCGAC---AAGGACGAGTGGCGCCTCTTC 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
DB 1091 CTATCAGCTCTCTGAGGCGCCAGCTTCTGCGCTGCGAGGCTGTGTGGAGACCATCTTT 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
DB 1151 CTGGTTTCCAGGCGCTGGATGCCAGGAGCTCCCGCGAGTTGCCCGCTGCCCGCAGCGC 1210
QY 391 TyrTyrGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
DB 1211 TACTGGCAATGCGGCGCTTCTTCTGGAGCTGTCTGGGAACACGCGCAGTGCCTTAC 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
DB 1271 GGGTGTCTCTCAAGACGAGCTGCCCGCTGCGAGTGC---GTACCCCGACGA 1321
QY 431 Leu-----Asn 432
DB 1322 GCCGCTGTCTGCCCGGAGAACCCAGGCTCTGTGCGCGCGCGCGCGAGGAGGAC 1381
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrGlnValTyr 452
DB 1382 ACACACCCCGCTGCGCTGTGTCAGCTCTCCGCGACACAGCGCCCTGGCAGGTGTAC 1441
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTyrGlyThrArgHis 472
DB 1442 GCCTTCTGCGGGCTGCTGCTGCGCGCTGCTGCCCGCGAGCTCTGGGGCTCCAGGAC 1501
QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
DB 1502 AACGAAACCGCGCTTCTCAGGAACACCAAGAGTTCATCTCCCTGGGAGGATGCCAAG 1561
QY 493 LeuSerLeuGlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrLeuArgSer 512
DB 1562 CTCCTGCTGCGAGGCTGACGTGGAAGTGGCGGAGTCTGCTGCTGCTGCTGCTGCTG 1621
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
DB 1622 AGCCACGGGCTGCTGCTGCTGCGCGCGAGACCGCTCTGCTGAGGAGATCTTGGGC 1681
QY 533 ThrPheLeuPheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552
DB 1682 AAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1741
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTyrSer 572
DB 1742 GTCAAGGAGACCGCTTCAAAAGAACAGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
DB 1802 AAGTTGAAAGCATTTGAATCAGACAGCATTTGAAGAGGTCAGCTGCGGAGCTGTG 1861
QY 593 GlnGluValArgHisHisGlnAspThrTyrTrpLeuAlaMetProLysCysArgLeuArg 612
DB 1862 GAAGCAGAGGTTCAGGAGCATCGGAAGACCGCGCGCTGCTGACGCTCCAGACTCCGC 1921
QY 613 PheLeuProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
DB 1922 TTCTATCCCAAGCTGACGGGCTGCGCGCGATTTGTGAACATGAGCTACGTCTGGAGGCC 1981
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
DB 1982 AGACGTTCCGACAGAAAGAGGCGCGCGCTCTCACCTCGAGGCTGAAGGACACTGTTC 2041
  
```

653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerValLeuGly 672  
2042 AGGTGCTCAACTACAGAGCGGGCGCGCGCTCTCTGGGCGCTCTGTGTGGG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValAlaLeuAspGln 692  
2102 CTGACCATATCCACAGGGGCTGGCGACCTTCTGTCTGTGTGTGGGCGCCAGACCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyValTyrAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGGTGGATGTACGGGGCGGTACACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGGCTACCGGAGGTCTACCGCAGCATCATCAAA---CCCAAGAACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValAlaValArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTGGGTATGCGGTGTCAGAGAGCGCGCCATCGGCACGTCCGCAAGCCCTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACCTTGCAGACCTCCAGCGGTACATCGACAGTGTGTGGCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGGATGCGTCTCATCGACAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValLysLys 812  
2453 GAGCCAGAGAGTGGCTCTTCGACGCTTCTCTACGCTTCATGTGCCACACCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTACGTCCAGTGCAGGGATCCCGAGGGCTCCATCTCTCCACG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGACACCTGTGCTGCGGACATGAGAACAGCTGTTCCGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCGCTTGTGGTGGATGATTTCTTGTGTGACACCTCACCTACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 CGCAAAACCTTCTCAGAGACCTTGTGCGAGGTGCTCCGTGATGGCTGCGTGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAAGACAGTGTGTGAATCTCCCTGTAGAACAGAGCGCCCTGGGTGGCAGCTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGGCCACCGCTATTTCCTGCTGTGGCGCTGCTGTGTGATACCGGACC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGAGGTGCGAGCGCTACTCCAGTATGCGCGACCTCCATCAGCCAGTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
2933 TTCAACCGCGCTTCAAGGCTGGGAGAACATGCTCGCAAACTTTTGGGTCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
2993 CTGAAGTGTACACCTGTTCTTGATTTGAGTGAACAGCTCCAGACGGTGTGCACC 3052  
993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
3053 AACATCTACAAGATCTCTCTGCTCAGCGGTACAGGTTTTCACGCATGTGTGTGACGCTC 3112  
1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032

3113 CCATTTCATCAGCAAGTTTGGAGAACCCACATTTTCTTGGCGCTCATCTCTGACAG 3172  
1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
3173 GCTCTCCTCTGTCTATCTCATCTCTGAAGCCAGAACGAGGGATGTCTGGGGCCCAAG 3232  
1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
3233 GGGCGCGCGCGCTCTCCCTCCGAGCGCGCTGCGAGTGTGCGCACCACCAAGCATTTCTG 3292  
1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
3293 CTCAGCTGACTCGACACCGTGTCACTACGTGCCACTCTCTGGGGTCACTCAGACAGCC 3352  
1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
3353 CAGACCGACTGAGTGGAGAGCTCCGGGGAGCGCTGACTGCCCTGGAGCGCGCAGCC 3412  
1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
3413 AACCCGCGCACTGCCCTCAGACTTCAAGACCATCTCTGGAC 3451

## RESULT 6

US-10-053-758-224  
; Sequence 224, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Lingner, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 224:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "htrp"

/note= "human telomerase reverse

transcriptase (htrp) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

3-10-053-758-224

3-09-042-460-2

(1-1122) x US-10-053-758-224 (1-4015)

Ligment Scores:

red. No.:

score: 0

Percent Similarity: 3505.00

Best Local Similarity: 72.94%

Very Match: 62.36%

Indels: 59.40%

Gaps: 15

Length: 4015

Matches: 719

Conservative: 122

Mismatch: 260

Indels: 52

Gaps: 13

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCGCGCGCTCCCGCTGCGAGCGTGGCTCCCTGCTGCGCAGCCACTACCGCGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGluProGluGlyArgArgLeuValGln 40  
 116 GTGCTGCGCGCTGGCCACCTTCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCTG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CCGGGGACCGCGCGCTTTCGGCGCGTGGTGGCCAGTGCCTGTGTGGCGCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GAGCGACGCTG 295  
 81 ValAlaArgValAlaGlnArgLeuCysGluArgAsnGluArgValLeuAlaPheGly 100  
 296 GTGGCGCGCGCTGTCAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCTT 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProMetAlaPheThrSerSerValArg 120  
 356 TTGCGCTGCTGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGCTACCTGCCCAACACCGTACCGACGACTGCGGGGGAGCGGGCGTGGGGCGTGG 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 476 CTGCGCGCGGTGGCGAGCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 535  
 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 536 CTGTGTGCTCCAGCTGGCGCTTACAGGTGTGCGGGCGCGCGCGCGCGCGCGCGCTG 595  
 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 596 GCCACTCAGCG 649  
 201 AsnPheThrAsnLeuArgPheLeuGlnIleLysSerSerSerArgGlnGluAlaPro 220  
 650 -----TGGCAACGCGCGCTGGAACCACTAGCGTCAGGAGCGCGCGG 688  
 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThr 240  
 689 GTCCCGCTGGCGCTGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 748

Qy 241 ValProSerAlaLysAlaArgCysTyrProValProAlaGluGluGlyPro--- 259  
 Db 749 CTGCCCTTCCCAAGAGGCGCGCGCTGGCGCTGCCCTGAGCGCGAGCGACCGCGTT 808  
 Qy 260 -----HisArgGlnValLeuProThrProSerGlyLysTrp--- 272  
 Db 809 GGGCAGGGTCTTGGGCCCAACCGCGCGAGGACCGTGGACCGAGTACCGGTGTTTGT 868  
 Qy 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 869 GTGGTGTCACTTGGCAGA-----CCGCCGGAAGAGCCACCTCTTTGGAG 913  
 Qy 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 914 GGTGCGCTCTTGGCAGCGCGCCACTCCACCCATCGTGGCGCGCAGCACCAACGCGGGC 973  
 Qy 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 974 CCCCACATCCACATCGCGCGCACCGTCCCTGGGACACGCTTGTCCCGCGGTGTACGCC 1033  
 Qy 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 1034 GAGACCAAGCATCTCTCTACTCTCTCAGGCGAC---AAGGAGCAGCTGCGCGCTCTTC 1090  
 Qy 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluLeuIlePhe 370  
 Db 1091 CTACTCAGCTCTCTGAGCGCCAGCTGACTGCGCTCGAGGCTCGTGAGACCATCTTT 1150  
 Qy 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 Db 1151 CTGGTTTCCAGGCGCTGGATGACGAGGACTCCCGCAGGTGCGCGCTGCGCGCGCGC 1210  
 Qy 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1211 TACTGGCAATCGCGCGCTCTTCTGGAGCTGTCTGGGAACCAACGCGCAGTGCCTTAC 1270  
 Qy 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 Db 1271 GGGGTGCTCTCAGACGACGCTGCGCGCTGCGAGTGGC-----GTACCCACGCA 1321  
 Qy 431 Leu-----Asn 432  
 Db 1322 GCCGTGCTGTGCCCGGAGAACCCAGGCTCTGTGGCGCGCCCGAGGAGGAGGAC 1381  
 Qy 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1382 ACAGACCCCGCTGCTGTGTCAGCTGCTCCCGCAGCAGCAGCGCGCTGCGAGTGTAC 1441  
 Qy 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1442 GCTTCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCGCGCTGCGCGC 1501  
 Qy 473 AsnGluArgPhePheLysAsnLeuLysPheLysPheLysPheLysPheLysPheLys 492  
 Db 1502 AACGAACGCGCTTCTCAGGAACACCAAGAAATTCATCTCCCTGGGGAAGCATCCCAAG 1561  
 Qy 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 Db 1562 CTCCTGCTGAGAGCTGACGTGGAGATGAGGTGCGGAGTGGCTTGGCTGGCTGGCAGG 1621  
 Qy 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
 Db 1622 AGCCACGGGTGGCTGTGTTCCGCGCGCAGACGCTCTGCTGAGGAGATCTCTGGCC 1681  
 Qy 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 Db 1682 AAGTCTCTGACCTGGCTGATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAT 1741  
 Qy 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePhePhePhePhePhePhePhePhe 572  
 Db 1742 GTCAGGAGACGAGTTTCAAAAGAACAGAGCTCTTTTCTACGGGAAGAGTGTCTGGAGC 1801  
 Qy 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592

1802 AAGTTGCAAGCATTTGATCAGACAGCACTTGAAGAGGGTGCAGCTGGGGAGCTGTGC 1861  
593 GlnGluGluValAlaArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAACGAGAGGTTCAGCAGCATCGGAGACCCAGGCGCCCTGCTGATGCTGACACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTrpSerMetGlyThr 632  
1922 TTATCCCAAGCTGACGGCTGCGGCGCATTTGTGAATGACATGACTACGTCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgLysGlnAlaHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCCGAGAGAAAGAGGCGGAGCGCTCACCTCGAGGGTGAAGGCACGTGTC 2041  
653 SerMetLeuAsnTrpGluArgThrLysHisPheProHisLeuMetGlySerValLeuGly 672  
2042 AGCGTGTCTCAACTACGAGCGGCGCGCGCCCTCTGCGGCGCTCTGTGCTGGGC 2101  
673 MetAsnAspIleTrpArgThrTrpArgAlaPheValLeuArgValAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGCGCTGCGGCACCTTGTGTGCTGTGCGGCGCCAGGACCG 2161  
693 ThrProArgMetTrpPheValLysAlaAspValThrGlyAlaTrpAspAlaIleProGln 712  
2162 CCGCTGAGCTGTACTTGTCAAGTGGATGTGACGGCGCGTACGACACCACTCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTrpCysIle 732  
2222 GACAGGCTCAGGAGGTTCATGCCAGCATCATCAA---CCCGAGACAGTACTCGTG 2278  
733 ArgGlnTrpAlaValAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTCGTATGCGGTGTCAGAGCGCGCCATCGGCACCTCGCGAAGCGCTTCAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTrpMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGTCTCTACTTGACAGACCTCAGCGGTACATGCGACAGTTCGTGGTCACTCCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCGCTCAGGAGTCCGCTGCTATCAGCAGAGCTCCTCCCTGAAT 2452  
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
2453 GAGCGCAGCAGTGGCTTCGAGCTTCTCAGCTTCATGCTCAGTTCATGCGCACACCGCGCGC 2512  
813 IleGlyAspArgCysTrpThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGCAAGTCTAGTCCAGTGCAGGGGATCCGCGAGGGCTCCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTCAGCCCTGTGTACGGGACATGAGAACAAAGCTGTTGCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTCGCTTGGTGGATGATTTCTGTTGGTGACACTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTrpGlyCysMetIleAsn 892  
2693 GCGAAACCTTCTCAGGACCTGCTCGAGGTGTCTCAGTATGGCTGCGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAGACAGTGGTGAATCTCCCTGTAGAGACGAGCGCTGGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTrpCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATGCGCGCCACCGGCTATTCCCTCGTGGCGCTGCTGCTGATACCCGAC 2872  
933 LeuGluValPheCysAspTrpSerGlyTrpAlaGlnThrSerIleLysThrSerLeuThr 952

Db 2873 CTGAGGTGCAGAGCGACTACTCCAGCTATGCCGAGCCTCCATCAGAGCCAGTCTCACC 2932  
QY 953 PheClnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuSerValLeuArg 972  
Db 2933 TTCACCGCGGCTTCAAGGTGGAGAAACATGCGTGCMAAATCTTTGGGGTCTTGGCG 2992  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2993 CTGAGGTGCACAGCTGTTTGGGATTTGAGGTGAACAGAGCTCCAGAGCGGTGTGCACC 3052  
QY 993 AsnIleTrpLysIlePheLeuGlnAlaTrpPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAGATCTCTCTGCTGCGGCGTACAGGTTTTCACGCATGTGTGTGAGCTC 3112  
QY 1013 PropheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTTCATCAGCAAGTTTGGAGAACCCCAATTTTCTGCGGTCTATCTCTGACAGC 3172  
QY 1033 AlaserCysCysTrpAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCCTCTCTACTCCATCTCTGAAAGCAAGACGAGGATGTGCTGGGGCCCAAG 3232  
QY 1053 GlySer-----PheProProGluAlaAlaHisTrpLeuCysTrpGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCAACCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTrpLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAGGTGACTCGACACCGTGTCACTACGTGCCACTCTCTGGGTCACTCAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGAGCTGAGTCGAGCTCCGCGGAGCGACGCTGACTGCCCTGAGGCGCGAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGCACTGCCCTCAGACTTCAGACATCTCTGGAC 3451

## RESULT 7

US-10-208-243-1  
; Sequence 1, Application US/10208243  
; Publication No. US20030044394A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaeta, Federico C.A.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune  
; FILE OF INVENTION: Response to a Telomerase Antigen  
; FILE REFERENCE: 015389-003500PC  
; CURRENT APPLICATION NUMBER: US/10/208,243  
; CURRENT FILING DATE: 2002-07-30  
; PRIOR APPLICATION NUMBER: US/09/675,321  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/112,006  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898  
; PRIOR FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patencin Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION: human telomerase reverse transcriptase (hTERT)  
US-10-208-243-1

Alignment Scores:  
Pred. No.: 0 Length: 4015  
Score: 3505.00 Matches: 719  
Percent Similarity: 72.94% Conservative: 122  
Best Local Similarity: 62.36% Mismatches: 260

Query Match: 59.40% Indels: 52  
B: 15 Gaps: 13  
S-09-042-460-2 (1-1122) x US-10-208-243-1 (1-4015)  
Y 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
b 56 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCACTACCGCGAG 115  
Y 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
b 116 GTGCTGCGCTGCGCAGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 175  
Y 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
b 176 CCGCGGAGCCCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 235  
Y 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerLeuLysGluLeu 80  
b 236 GAGCGAGCGCGCGCGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 295  
Y 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
b 296 GTGGCGCGAGTCTGCGAGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 355  
Y 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
b 356 TTCGCGCTGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415  
Y 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
b 416 AGCTACCTCCCAACACGCGTGACCGACACCTGCGGGGAGCGGGCGGTGGGGCTG 475  
Y 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
b 476 CTGCGCGCGCTGGCGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 535  
Y 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
b 536 CTGCTGCTCCAGCTGCGCTACCGAGTGTGCGCGCGCGCGCTGCTACCGAGCT 595  
Y 181 ThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
b 596 GCACCTCAGCG 649  
Y 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220  
b 650 -----TGCAGCGCGCGCTGGAACCATAGCGTCAAGGAGCGCGCG 698  
Y 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerSer 240  
b 689 GTCCCTCTGGCTGCGACCGCGCGCTGCGAGGCGCGCGCGCGCGCGCGCGCGCG 748  
Y 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259  
b 749 CTGCGCTGCGCCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808  
Y 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
b 809 GGCAGCGGTCTCG 868  
Y 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
b 869 GTGGTGTACCTGCCAGA-----CCCGCGGAAGAGCGCACCTCTTTGGAG 913  
Y 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
b 914 GGTGCGCTCTGTGCGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973  
Y 312 SerThrSerLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---Phe 330  
b 974 CCCCGATCCACATCG 1033

331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
1034 GAGACCAAGCAGCTTCTCTACTCTCTCAGCGGAC---AAGAGCAGCTGGCGCCCTCTTC 1090  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
1091 CTACTCAGCTCTCTGAGGCCAGCTGCTGCGCTGCGAGGCTGCTGGAGACCATCTTT 1150  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
1151 CTGGTTTCCAGGCGCTGAGGCCAGGACTCCCGCAGGTGCTCCCGCTGCCAGCGC 1210  
391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
1211 TACTGGCAATGCGGCCCTGTTCTGGAGCTGCTGGGAACCAAGCGAGTGGCCCTAC 1270  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
1271 GGGTGTCTCTCAAGAGCAGCTCCCGCTGCGAGTGGC-----GTACCCCGAGCA 1321  
431 Leu-----Asn 432  
1322 CCCGCTGTGTCGCGGAGAGCCCGAGGCTCTGTGCGCGCCCGCGAGGAGGAC 1381  
433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
1382 ACAGACCCCGCTGCGCTGTGTCAGCTGCTCCGCGCAGCAGCAGCCCGCTGGCAGTGTAC 1441  
453 GlyPheLeuArgAlaCysLeuValValSerAlaSerLeuTrpGlyThrArgHis 472  
1442 GCTTGTGCGGCTGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCT 1501  
473 AsnGluArgArgPhePheLeuLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
1502 AACCAACGCGCTTCTCAGCAACACACAGAGATTCTCTCCCTGGGGAAGCATGCCAAG 1561  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
1562 CTCGCTGCGAGGAGCTGACGTGGAAGATGAGCTGCGGAGCTGCGCTGGCTGGCGCAG 1621  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
1622 AGCCAGCGGTGCTGCTGTTCCGCGCGCAGACACCGTCTGCTGAGGAGATCTCTGGCC 1681  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552  
1682 AAGTCTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1741  
553 IleThrGluSerThrPheGlnLysAsnArgPhePheTyrArgLysSerValTrpSer 572  
1742 GTCAGGAGACCATGTTTCAAGAGACAGGCTCTTTTCTACCGGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAACCATGGAATCAGACAGCTTGAAGAGGCTGAGCTGCGGAGCTGCTGCG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGTCAAGAGCATCGGAAGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCATCCCAAGCTGAGCGGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCGCG 1981  
633 ArgAlaLeuGlyArgArgGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGACGTTCCGAGAGAAAGAGGCGCGAGCTCTACCTCGAGGGTGAAGGACCTGTTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGTCTCAACTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuArgGln 692





LOCATION: 56..3454

OTHER INFORMATION: /product= "htrr"  
 /notes "human telomerase reverse  
 transcriptase (htrr) catalytic protein  
 component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

0-054-295-224

ment Scores:

L. No.: 0 Length: 4015  
 e: 3505.00 Matches: 719  
 ent Similarity: 72.94% Conservative: 122  
 Local Similarity: 62.36% Mismatches: 260  
 y Match: 59.40% Indels: 52  
 15 Gaps: 13

19-042-460-2 (1-1122) x US-10-054-295-224 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCCCGCGCTCCCGCTGCCAGCGCTGGCTCCCTGCTGGCAGCCACTACCGCGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40  
 116 GTGCTCCGCTGCCACGCTTCGTGGCGGCTGGGGCCCGCAGGCGTGGCGCTGGTCAG 175  
 41 ProGlyAspProGlyTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60  
 176 CGGGGGAGCCCGCGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGGCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuGluLeu 80  
 236 GAGCGACGCGCGCCCGCCCGCTCTTCGCGCAGGTGCTCGCTGAAGAGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGGCCGAGTGTGTCAGAGGCTGTGGAGCGCGCGCGAGAGAGCTGCTGGCTTCGGC 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGTGACGGGCGCGCGGGGCCCCCGCGAGGCTTCAACACGAGCTGGCGC 415  
 121 SerTyrLeuProAsnThrValLeuGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGTACCTGTCACACAGCTGACCGACCTGCGGGGAGCGGGGCTGGGGCTGGCTG 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 476 CTGGCGCGCTGGCGACGACGCTGTGTCTACCTGCTGGCAGCTGGCGCTCTTTGTG 535  
 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 536 CTGTGGCTCCAGCTGCGCTACCAAGGTGTGGGGCGCGCGCTGTACAGCTGGCGCT 595  
 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 596 GCGACTCAGCG 649  
 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
 650 -----TGCAGACCGCGCTGGACCATAGCGTCCAGGAGCGCGCGG 688  
 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
 689 GTCCCCCTGGCGCTGCCAGCGCGCGGTGGAGAGCGCGCGCGCGCGCGCGCGCGCG 748  
 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259  
 749 CTGCGGTGCCAGAGCG 808  
 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp 272  
 809 GGGCAGGGGTCTGGGCG 868

QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 DB 869 GTGTGTCTCCTGCGCAG-----CCGCGCGAAGAGACCCACCTCTTTGGAG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 DB 914 GGTGGCTCTCTGGCAGCGCGCCTCCACCCACTCGTGGCGCGCAGCACCAGCGGGC 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 DB 974 CCCCACATCCACATCGCGGCCACCGCTCCCTGGGACACGCCCTTGTCCTCCCGGTGTACGC 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 DB 1034 GAGACCAAGCACTTCTCTACTCTCAGGCGAC---AAGGAGCAGCTGGCGGCTCTCTTC 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370  
 DB 1091 CTACTCAGCTCTCTGAGGCCAGCTGACTGGCGCTCGGAGGCTCGTGAGACCATCTTT 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390  
 DB 1151 CTGGGTTCCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCTGCCCGCAGCGC 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
 DB 1211 TACTGGCAATGCGGCCCTGTTTCTGGAGCTGTTGGGAACCAACGCGCAGTGGCCCTAC 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
 DB 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTCGAGCTGCG-----GTCACCCAGCA 1321  
 QY 431 Leu-----Asn 432  
 DB 1322 GCCGTGTCTGCGCGCGGAGAACCCCGAGGCTCTGTGTGGCGCCCGCAGGAGGAGC 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 DB 1382 ACAGACCCCGCTCGCTGTGTGCTGCTCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1441  
 QY 453 GlyPheLeuArgAlaCysLeuValValSerAlaSerLeuTrpGlyThrArgHis 472  
 DB 1442 GCTTCGTGGCGGCTGCTGCTGGCGGCTGGTGGCGCCCGCCTCTGGGCTCCAGGCAC 1501  
 QY 473 AsnGluArgArgPhePheLeuAsnLeuLysPheLeuSerLeuGlyLysTyrGlyLys 492  
 DB 1502 AACGAACCGCTTCTCAGGAACACCAAGAAAGTTCTCTCCCTGGGGAAGCATGCCAAG 1561  
 QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 DB 1562 CTCTCGTGCAGAGCTGACCTGGAAGATGAGCGTGGGAGCTGGCTGGCTGGCGGAGG 1621  
 QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGileLeuAla 532  
 DB 1622 AGCCAGGCGTGGCTGTGTTCCGCGCGCAGACGCGCTCTGCGTGAGAGATCCTGCC 1681  
 QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 DB 1682 AAGTTCCTGCACTGGCTGATGAGTGTGTACGTGTGTGAGCTGCTCAGGTCTTTTAT 1741  
 QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
 DB 1742 GTCACGAGACCATGTTTCAAAAGAACAGGCTCTTTTCTACCGGAAGAGTGTCTGAGC 1801  
 QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
 DB 1802 AAGTTGCAAGCATTTGGAATCAGACGACTTGAAGAGGGGTGAGCTGCGGAGCTGTGC 1861  
 QY 593 GlnGluValValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
 DB 1862 GAAGCAGAGCTCAGCAGCATCGGAAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1921

613 PheileProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCCTGACGGGCTGCGCGGATGTGAACATGGACTACGTCGTTGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAACGTTCCGACAGAAAGAGGCGCGGAGGCTCTCACTCGAGGGTGAAGGCACTGTT 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCAACTACGAGCGGCGCGCGCGCGCTCTCTGGCGCTCTGTGCTGGCG 2101  
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValAlaLeuAspGln 692  
2102 CTGGAGCATATCCAGAGGCTGGCGGCACTTGTGCTGCTGCTGCGGCGCGGAGCCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCGCTGAGCTGACTTGTCAAGGTGATGTACGGCGCGGTACGACACCATCCCGCAG 2221  
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
2222 GACAGCTCACGAGCTCATCCCGACATCATCAA---CCCGACACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCAGAAAGCGCCCATGGGACGTCGCCAAGGCGTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGCTCTACCTTGACAGACCTCCAGCGGTACATCGACAGTTCGTGGGTCACTCGAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCGCTGAGGATGCGGTGTCATCGAGAGAGCTCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGGCGCAGAGTGGCTCTTCGACGCTCTCTTACGCTTCATGTGCCACCAAGCGCGTCCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTTACGTCAGTCCAGGGGATCCCGGAGGCTCCATCTCTCCAG 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGGCTGTGCTACGCGGACATGAGAAAGAGTGTTCGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGCTGCTCTGCGTTGGTGGATGATTTCTTTGGTGACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCAGAAACCTTCTCAGGACCTGCTGCGAGGTGCTCTGAGTATGGTGGTGGTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTCGGGAAGACAGTGTGATCTTCCCTGTGAGAGACGAGGCGCTGGGGGCGGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCGCGCCACGCGCTATTCCCTGTGCGGCTGCTGTGATACCCGAGC 2872  
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
2873 CTGGAGGTGAGAGGCACTACTCCAGCTATGCCGAGCTCCATCAGAGCGAGTCTCACC 2932  
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
2933 TTCACCGCGGCTTCAAGGCTGGAGGACATGCTCGCAACTCTTTGGGCTTGGCG 2992  
973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992

Db 2993 CTGAAGTGTACAGCCTGTTTCTGATTTGCAGTGAACAGCCTCCAGCGTGTGCACC 3052  
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGCTGCGAGCGGTACAGGTTTTCACGATGTGTGTGCAGCTC 3112  
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTCACTCAGCAAGTTTGGAGAACCCACATTTTCTCGCGCTCATCTCTGCACAG 3172  
Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCCTCTGCTACTCCATCTGAAAGCCAGAACGCGGAGTGTGCTGGGGGCGCAAG 3232  
Qy 1053 GlySer-----PheProProGluAlaAlaHisIleTyrLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GCGCGCGCGCGCGCTCTGCGCTCGAGGCGGTGCTGAGTGGCTGTGCCACCAAGCATTTCTGT 3292  
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGTGTACTCGACACCGGTGTCACTAGTCCACTCTCTGGGGTCACTCAGAGCAGCC 3352  
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAlaAla 1109  
Db 3353 CAGACGCGAGCTGAGTCTGAAAGCTCCCGGGGACGCGTGAAGTGTGCTGGAGCGCGAGCC 3412  
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCGGACATGCGCTCAGACTTCAGACCATTCCTGGAC 3451

## RESULT 9

US-10-054-611-224  
Sequence 224, Application US/10054611  
Publication No. US20030059787A1  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030059787A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,611  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 224:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"

/note= "human telomerase reverse

transcriptase (hTRT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 224:

-10-054-611-224

Alignment Scores:

Seq. No.:	Length:	Matches:	Mismatches:	Indels:	Gaps:
0	3505.00	719	122	260	13
1	3505.00	719	122	260	13
2	72.94%	122	260	13	
3	62.36%	260	13		
4	59.40%	13			

-10-054-460-2 (1-1122) x US-10-054-611-224 (1-4015)

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 56 ATGCGCGCGCTCTCCCTGCTCCAGCGCGTGGCTCTCTCTGCGGAGCACTACCGCGAG 115  
 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluArgArgLeuValGln 40  
 116 GTGCTGCGCTGGCCACGTTCTGTCGCGCGCTGGGCGCCAGCGGTGGCGGTGGCGAG 175  
 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethIleTrp 60  
 176 CGCGGGACCGCGCGCTTTCGCGCGCTGGTGGCCAGTGTCTGTGTGGTGGCTGG 235  
 61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 236 GACGACGCGCGCGCGCGCGCTCTCTCCGCGAGGTGTCTGCTGCTGAAGAGCTG 295  
 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 296 GTGCGCGAGTGTGCAGAGGCTGTGCGAGCGCGCGCGGCGGAGAACGCTGTGCGCTTCGCG 355  
 101 PheGluLeuLeuAsnGluAlaArgGlyProProMetAlaPheThrSerSerValArg 120  
 356 TTCGCGCTGTGGACGGGCG 415  
 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 416 AGCTACTCTGCCAACACCGGTGACGACCGCACTGCGGGGAGCGGGCGGTGGGTGGT 475  
 141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 476 CTGCGCGCGTGGGCGGACGAGTGTGGTTCACCTGTCGACGCTGCGCGCTCTTTGTG 535  
 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
 536 CTGTGTGCTCCAGCTGCGCTACAGGTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 595  
 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
 596 GCCACTAGCG 649  
 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220

Db 650 -----TGCAGACGGCGCTGGAAACCATAGCGTCAGGAGGCGCGG 688  
 QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240  
 Db 689 GTCCCGCTGGCGCTGCCAGCGCGCGGTGCGAGAGGCGCGGGGCGAGTGCACCGAAGT 748  
 QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259  
 Db 749 CTGCGGTTGCCAAGAGCGCCAGCGCGTGGCGCTGCCCTGAGCGGAGCGGCGCGGTT 808  
 QY 260 -----HisArgGlnValLeuProThrProThrProSerGlyLysSerTrp--- 272  
 Db 809 GGGCAGGCGTCTCTGCGGCCACCGCGGCGAGCAGCGTGGACCGAGTACCGTGTCTCTGT 868  
 QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292  
 Db 869 GTGGTGTACCTGCCAGA-----CCCGCCGAAGAAGCCACCTCTTTGGAG 913  
 QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311  
 Db 914 GGTGCGCTCTTGGCAGCGCGCACTCCACCCATCTGTCGCGCGCCAGCACACCGCGGCG 973  
 QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
 Db 974 CCCCCATCCACATCGCGGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTGACGCC 1033  
 QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350  
 Db 1034 GAGACACAGCACTTCTCTACTCTCTCAGCGGAC---AAGAGAGCAGCTGCGGCGCTCTTC 1090  
 QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
 Db 1091 CTACTCAGCTCTCTGAGGCCAGCGCTACTGCGCTCGAGGCTCGTGGAGCATCTTT 1150  
 QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
 Db 1151 CTGGGTTCAGCGCCCTGGATCCAGGAGCTCCCGCAGGTGCCCCCGCTGCCCGACGC 1210  
 QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410  
 Db 1211 TACTGCAATCGCGCGCTGTTTCTGGAGCTGTGGAAACACACGCGCGAGTGCCTTAC 1270  
 QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430  
 Db 1271 GGGGTGCTCTCAAGACGCACTGCGCGCTGCGAGTGG-----GTACCCCGACGA 1321  
 QY 431 Leu-----Asn 432  
 Db 1322 GCGGTGTCTGTCCCGGAGAGCGCCCGAGGCTCTGTGGCGCGCCCGAGGAGGAGAC 1381  
 QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
 Db 1382 ACAGACCGCGCTGCGCTGTGTGTCAGCTGCTCGCGCAGCAGCAGCGCGCTGGCAGGTGTAC 1441  
 QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
 Db 1442 GAGTTCGTGGCGCTGCTGCGCGGTGTGTGCCCCCAGCGCTCTGGGGTCCAGGCAC 1501  
 QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
 Db 1502 AACGAAACGCGCTTCTCAGAAACACCAAGAGTTTCATCTCCCTGGGGAACATGCCCAAG 1561  
 QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
 Db 1562 CTCCTGCTGAGAGGTGACGTGGAGATGACGTGCGGAGCTGGCGTGGCTGGCGAGG 1621  
 QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
 Db 1622 AGCCAGGCGTGGTGTGTCTCCGCGCGCAGACACCGCTCTGCGTGAAGAGATCTCTGCC 1681  
 QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
 Db 1682 AAGTTCCTGACCTGGCTGTGATGTGTACGTGTGCTGAGCTGCTCAGGTCTTCTTTTAT 1741

553 IleThrGluSerThrPheGlnIysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
1742 GTTACGGAGACCAAGCTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGlnArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGCAATTTGAATCAGACAGCACTTTGAAGAGGTGACAGTCCGGAGCTGTGG 1861  
593 GlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProLysCysArgLeuArg 612  
1862 GAAGCAGAGGTGAGCAGCATCGGAAGACCAAGCCCGCTCTGAGCTCCAGACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProLysValAlaMetSerTyrSerMetGlyThr 632  
1922 TTTCATCCCAAGCTTGACGGCTGCGGCCGATTTGTGAACATGACTACGTCTGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
1982 AGAAGCTTCCGAGAGAAAGAGGCGCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 ACGGTGCTCACTACGAGCGGGCGCGGCCCGCTCTCTGGCGGCTCTGTGCTGGCG 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGACGATATCCACAGGCGCTGGCGACCTTCTGTGTGGTGTGGCGGCCAGGCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGGCTGAGCTGATCTTGTCAAGTGTGATGACGGCGCGCTACGACACATCCCGCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGlnSerThrTyrCysIle 732  
2222 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGACACGATCTGGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGTGTGTCAGAAAGCGCGCCATGGGCACGCTCCGCAAGGCTTCAAGAGC 2338  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CACGTCTCTACCTTGACAGACTCCAGCGCTACATGCGACAGTTCGGTCTCACCTGCAG 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATGCCGTCTCATCGACGAGAGCTCTCCCTGAAT 2452  
793 SerSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValLys 812  
2453 GAGGCGAGAGTGGCTCTTCAGCTCTTCTACGCTTCATGTCGTCACCAAGCGGTGGCG 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGTCTCTACGTCCAGTCCAGCGGATCCCGCAGGCTCCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGCTGTGATCGGCGACATGAGAACCAAGCTGTTTGGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuArgPheValAspAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTCTCTCGGTTTGGTGGATGATTTCTTTGTTGTGACACCTCACCTCACCCAC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTCTCTCAGGACCTGTGTCGAGGTGTCTCCAGTATGCTGCTGCTGCTGAAC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGGGGAAGACAGTGTGTGAATCTTCTCTGTAGAACGAGGGCCCTGGGTGGACAGGCTTTT 2812

QY 913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
Db 2813 GTTCAGATGCGGCGCCACGCGCTATTCCCTGTGTGGCTGTCTGTGTGATACCCGAGCC 2872  
QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
Db 2873 CTGGAGGTGACAGAGCACTACTCCAGCTATGCCCGACCTCCATCAGACCGAGTCTCACC 2932  
QY 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
Db 2933 TTCAACCGCGCTTCAAGGCTGGGAGAACATCGTCGCAAACTCTTTGGGGTCTTGGCG 2992  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
Db 2993 CTGAAGTGTACAGCTGTCTTCTGGATTTGAGGTGAAACAGCTCCAGCGGTGTGCACC 3052  
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
Db 3053 AACATCTACAAGATCTCTCTGTCTGAGGCGGTACAGGTTTTCAGGCATGTGTGTGCAGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
Db 3113 CCATTTCTATCAGCAAGTTTGGAAAGAACCCACACATTTTCTGGCGTCTATCTGTACAGC 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
Db 3173 GCCTCCCTCTGTACTCCATCTCTGAAAGCCAGAACGAGGATGTGCTGGGGCCAAAG 3232  
QY 1053 GlySer-----PheProGluAlaAlaHisTrpLeuCysTyrGlnAlaPheLeu 1069  
Db 3233 GGCGCCCGCGGCTCTGCGCTCCGAGGCGGTGAGTGTGTCGCCCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
Db 3293 CTCAAGCTGACTGCGACACCGTGTCACTAGTCCCACTCTGGGGTCACTCAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
Db 3353 CAGACGAGCTGAGTCCGAAGCTCCCGGGAGACGACCTGACCTGCGCTGGAGGCGGAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
Db 3413 AACCCGCACTGCGCTTCAGACTTCAGACCATCTCTGCAC 3451

## RESULT 10

US-10-105-963-1  
; Sequence 1, Application US/10105963  
; Publication No. US20030068818A1  
; GENERAL INFORMATION:  
; APPLICANT: Geron Corporation  
; APPLICANT: Denning, Chris  
; APPLICANT: Schiff, J. Michael  
; TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
; TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
; TITLE OF INVENTION: Recombination  
; FILE REFERENCE: 731/002  
; CURRENT APPLICATION NUMBER: US/10/105,963  
; CURRENT FILING DATE: 2002-03-21  
; PRIOR APPLICATION NUMBER: US 60/277,811  
; PRIOR FILING DATE: 2001-03-21  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 4015  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (56)..(3454)  
; OTHER INFORMATION:  
US-10-105-963-1

## Alignment Scores:

```

ed. No.:      0      Length:      4015
ore:          3505.00  Matches:      719
cent Similarity: 72.94%  Conservative: 122
st Local Similarity: 62.36%  Mismatches: 260
ery Match:      59.40%  Indels: 52
:               15      Gaps: 13

-09-042-460-2 (1-1122) x US-10-105-963-1 (1-4015)

1  MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56  ATCCCGGGCGCTCCCGCTCCCGAGCGCTGCTCCCTGCTGGCGAGCCACTACCGCGAG 115
21  ValTrrProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGCTCCGCTGGCCACAGTTCGTGGCGGCCCTGGGGGCCCGAGGCTGGCGCTGGTCG 175
41  ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrr 60
176 CGCGGGGACCCGCGGCTTTCGCGCGCTGGTGGCCAGTGCCTGGTGGTGGCTGG 235
61  GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuGlyGluLeu 80
236 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 295
81  ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGCGCGAGTGTCTCAGAGCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrrMetLeuLeu 140
416 AGTACCTGCTCCCAACAGGTGACCGACGACCTGCGGGGAGCGGGGCGGTGGCG 475
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTrrLeu 160
476 CTGCGCGCGTGGCGACGAGTGTGTTACCTGCTGGCACGCTGCGCGCTCTTTGTG 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGCTGCTCCCACTGCTGCTACCAAGGTGTGCGGCGCGCGCGCGCGCGCGCG 595
181 ThrThrAspIleTrrProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCNACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
650 -----TGCAGACGGGCTGCAACCATAGCTGACGAGGAGCGCGG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240
689 GTCCCTCTGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
749 CTGCGGTTCGCCAAGAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrr 272
809 GGGCAGGGGTCTCGGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
869 GTGTGTGTACCTGCCAGA-----CCGCGCGGAAGAACCCACCTCTTTGGAG 913
293 GlyLysValSerAspLeuSerLeuSer-----GlySerValCysCysLysHisLysProSer 311
914 GGTGCGCTCTCTGGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973

```

```

QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro----PheIle 330
Db 974 CCCCACATCCATCGCGGCCACACACCGCTTGTCCCGCGGTGACGCC 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACTTCTCTACTCTCTAGGGGAC---AAGAGGAGCTGGCGCTCTCTTC 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
Db 1091 CTACTCAGCTCTCTGAGGCCCGACCTGACTGGGCTCGGAGGCTCTGGAGACCATCTTT 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1151 CTGGGTTCCAGGCTTGGATGCCAGGACTCCCGCAGGTGGCCGCTCTGCCCGAGCG 1210
QY 391 TyrTrrPginMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
Db 1211 TACTGGCAATGCGGCCCTTCTTCTGGAGCTGCTTGGGAACACACGCGAGTGCCTAC 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
Db 1271 GGGGTGCTCTCAAGACGCACTGCCCGCTGCGAGCTGCG-----GTCACCCACGA 1321
QY 431 Leu-----Asn 432
Db 1322 GCGCGTGTCTGCGCGGAGAACCCCGAGGCTCTGTGGCGCCCGCGAGGAGGAC 1381
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrrPginValTyr 452
Db 1382 ACAGACCCCGCTGCGCTGTGTCAGCTCTCGCGCAGCACAGACGCGCTGGCAGGTGAC 1441
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrrGlyThrArgHis 472
Db 1442 GCTTGTGGGGCTGCTGCGCGGCTGGTGGCCCGCGGCTCTGGGCTCCAGGAC 1501
QY 473 AsnGluArgArgPhePheLysAsnLeuLysLysPheLysLeuGlyLysTyrGlyLys 492
Db 1502 AACGAACCGCTCTCTCAGGAACACCAAGAGTTCTCTCCCTGGGGAAGCATGCCAAG 1561
QY 493 LeuSerLeuGlnGluLeuMetTrrLysMetLysValGluAspCysHisTrrLeuArgSer 512
Db 1562 CTCTCGCTGCGAGAGCTGACGTGAAGATGAGCTGGGAGCTGCGCTTGGCTGGCAGG 1621
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
Db 1622 AGCCAGGGGTGGCTGTGTTCCGCGCGCAGACCGCTCTGGCTGAGGAGATCCTGGCC 1681
QY 533 ThrPheLeuPheTrrLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyr 552
Db 1682 AAGTTCTCTGCACTGGCTGATGAGTGTGTGCTGCTGAGCTGCTCAGTCTTCTTTAT 1741
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTrrSer 572
Db 1742 GTACGAGAGACCGCTTCTCAAGACACAGGCTCTTTCTACCGAAGAGTGTCTGGAGC 1801
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
Db 1802 AAGTTGCAAGCATTTGGAATCAGACAGCACCTTGAAGGGGTGCACTGCGGGAGCTGCG 1861
QY 593 GlnGluGluValArgHisHisGlnAspThrTrrLeuAlaMetProLleCysArgLeuArg 612
Db 1862 GAAGCAGAGGTTCAGCAGCATCGGGAAGCCAGGCGCGCGCTGTGAGCTCCAGACTCCGC 1921
QY 613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
Db 1922 TTCAATCCCAAGCCCTGACGGCTGCGGCGCGATTGTGAACATGGAACACTCGCTGGGAGCC 1981
QY 633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
Db 1982 AGAACGTTCGCGAGAAAGAGGCGCGAGCGCTCTCACCTCGAGGGGTGAAGGCACTGTTTC 2041

```





TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product="hTERT"

/note="human telomerase reverse

transcriptase (hTERT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

-10-044-692-1

Alignment Scores:

ed. No.:	0	Length:	4015
ore:	3505.00	Matches:	719
cent Similarity:	72.94%	Conservative:	122
st Local Similarity:	62.36%	Mismatches:	260
ery Match:	59.40%	Indels:	52
	15	Gaps:	13

-09-042-460-2 (1-1122) X US-10-044-692-1 (1-4015)

```

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGCGCGCGCTCCCGCTGCGAGCGTGCCTCTCTGCTGCGAGCCACCTACCGCGAG 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGTCGCTGCGCAGCTTCGTCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCGCTGCG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
176 CGCGGGGACCGCGGCTTTCGCGCGCTGCTGTCGCGCGCTGCTGTCGCGCGCTGCTG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuLeu 80
236 GACGACGCGCGCGCGCGCGCGCTCTCTCTGCGCGCGCTCTCTGCTGAGGAGCTG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGCGCGCGAGTGTGCAGAGCTGTGCGAGCGCGCGCGAGCGCGCGCGCTTTCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGGACGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTTCACACCGCTG 415
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
416 AGCTACCTGCGCCACACGCTGACCGACGCTGCGGGGCGCGGGGCGTGGGGCTGCTG 475
141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
476 CTGCGCGCGCTGGCGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
536 CTGTGTGCTCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
596 GCCACTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 649
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGluAlaPro 220
650 -----TGGGAACGGCGCTGGAACCATAGCTCAGGAGGCGCGG 688
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240

```

```

Db 689 GTCCCGCTGGCGCTGCCAGCCCGCGGTCGAGGAGCGCGCGGCGAGTCCACCGCAAGT 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
Db 749 CTGCGCTTGGCCAAAGAGAGCCCGCGGCTGCGCTGCGCTGAGCGGAGCGGCGCGCTT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
Db 809 GGGCAGGGGTCTCTGGGCCACCCCGGCGAGCGGTGGACCGGTGACCGTGGTTCCTGT 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
Db 869 GTGGGTGTCACTGTCAGCA-----CCGCGCGAAGAACCGACCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 914 GGTGGCTCTCTGGCAGCGCCACTCCACCCATCCGTGGCGCGCGCGACACCGCGGCG 973
QY 312 SerThrSerLeuLeuSerProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 974 CCCCATCCATCGCGGCCACCGCTCCCTGGGACACGCTTGTCCCGCGGTAGCGCC 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGGAC---AAGGAGCAGCTGCGCGCTCTTC 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
Db 1091 CTACTCAGCTCTCTGAGCGCCAGCGCTGCTGCGCTCGAGGCTCTGTGGAGACCACTTT 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1151 CTGGGTTCAGCGCCCTGATCCAGGAGCTCCCGCAGGTTCGCCCGCTGCCAGCGC 1210
QY 391 TyrTrpGlnMetArgProPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
Db 1211 TACTGGCAATGCGCGCTGTTCTGGAGCTGCTTGGGAACACCGCGCAGTGGCCCTAC 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430
Db 1271 GGGGTGCTCTCAAGACGCACTGCCGCGTGGAGCTGCG-----GTACCCCGACGA 1321
QY 431 Leu-----Asn 432
Db 1322 GCGGTGTCTGTGCCCGGAGAGCCCGAGGCTCTGTGGCGCGCGCGCGAGAGGAGAG 1381
QY 433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
Db 1382 ACAGACCGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1441
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
Db 1442 GGTTCGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1501
QY 473 AsnGluArgArgPhePheLeuAsnLeuLysLysPheLeuLeuLysLysLysLys 492
Db 1502 AACGACCGCGTCTCTCAGGAACACCAAGAGTTCATCTCCCTGGGAGACATGCCAAG 1561
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
Db 1562 CTCTGCTGCGAGGCTGACGTGGAAGATGAGCGTGGGAGCTGCGCTTGGCTGCGCAGG 1621
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532
Db 1622 AGCCAGGGGTGGTGTGTTCCGCGCGCAGAGCACCGTCTGCGTGAGGAGATCTTGCC 1681
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuLeuLeuLeuLeu 552
Db 1682 AAGTTCCTCAGCTGCTGATGAGTGTGACGTCTGCGAGCTGCTCAGGTCTTCTTTAT 1741
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572

```

1742 GTACGGAGACACAGCTTTCAAAAGNACAGGCTCTTTTCTACCGAAGAGCTGTCTGGAGC 1801  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
1802 AAGTTGCAAGACATTTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCGGAGCTGTGC 1861  
593 GlnGluGluValArgHisIleGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
1862 GAAGCAGAGGTGAGCAGCATCTGGAGAGCCAGGCCCGCTCTGTGAGCTCCACACTCCGC 1921  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
1922 TTCAATCCCAAGCCCTGAGCGCTGCGGCCGATGTGAACATGAGTACGTACGTGCGGGAGCC 1981  
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLeuThrLeuPhe 652  
1982 AGAAGCTTCCGAGAGAAAGAGGCCGCGCTCTCACCTCGAGGGTGAAGGCACTGTTC 2041  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
2042 AGCGTGCTCACTACAGCGGGCGCGGCCCGCTCTCTGCGGCGCTCTGTGTGGGC 2101  
673 MetAsnAspIleTyrArgThrTrpArgAlaPheValLeuArgValArgAlaLeuAspGln 692  
2102 CTGGAGATATCCACAGGCGCTGCGGCACCTCTCTGTGTGGTGTGGCGGCCACGAGCCCG 2161  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
2162 CGCCTGTAGCTGTACTTTGTCAAGTGTGTGACGGGCGGTACGACACCATCCCCAG 2221  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisGlnSerThrTyrCysIle 732  
2222 GACAGGCTCACGGAGTCTATCGCCAGCATCATCAAA---CCCAAGAACAGTACTGCGTG 2278  
733 ArgGlnTyrAlaValArgArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
2279 CGTGGTATGCGGTGTCCAGAGGCCCGCCATGGCAGCTCCCAAGGCTTCAAGAGC 2338  
753 GlnValThrLysSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
2339 CAGCTCTACCTTGACAGACCTCCAGCGGTACATGCGACAGTGTGCGGTCACTCCAGC 2398  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSerMetAsn 792  
2399 GAGACC-----AGCCCGCTGAGGATCCGCTGTCTATCGAGCAGAGCTCCTCCCTGAAT 2452  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
2453 GAGCGCAGCTGGCTCTTCGAGCTCTTCTACGCTTCATGTGCCACCGCGTGGC 2512  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832  
2513 ATCAGGGGCAAGCTCTAGCTCCAGTCCAGGGGATCCCGCAGGGCTCCATCTCTCCAGC 2572  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852  
2573 CTGCTCTGAGGCTGTCTACGGCGCATGTGAGAACAGCTGTGTGGGGATTCGGCGG 2632  
853 AspGlyLeuLeuLeuArgPheValAspPheLeuLeuValThrProHisLeuAspGln 872  
2633 GACGGGCTCTCTCGTGTGTGTGATGATTTCTGTGTGTGACACCTCACTCACTCCAGC 2692  
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892  
2693 GCGAAACCTCTCTCAGGACCTGTGTGAGGTGCTCTGATGATGCTGTGTGTGAGC 2752  
893 LeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaPro 912  
2753 TTGCGGAGACAGTGTGTGACTTCTCTGTAGAACGAGGCGCTGGGTGGCAGCGCTTTT 2812  
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932  
2813 GTTCAGATCCCGGCCACCGCTATTTCCTCTGTGTGGTGGCGCTGTGTGTGATACCGGACC 2872

QY 933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952  
DB 2873 CTGAGGTGAGAGCGGACTACTCCAGTATGCCGAGCCTCCATCAGACCCAGTCTCACC 2932  
QY 953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972  
DB 2933 TTCAACCGCGCTTCAAGGCTGGAGGACATGCGTGGCAAACTCTTTGGGTCTTGGCG 2992  
QY 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992  
DB 2993 CTGAAGTGTACAGCCTGTCTTCTGATTTGAGGTGACAGCTCCAGACGGTGTGCACC 3052  
QY 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012  
DB 3053 AACATCTACAGATCTCTCTGCTGCGAGGCTACAGTTTCACGCATGTGTGTGCAGCTC 3112  
QY 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032  
DB 3113 CCATTTCAATCAGCAAGTTTGGAAAGAACCCACATTTTCTGCGCGCTCATCTCTGACAGC 3172  
QY 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052  
DB 3173 GCCTCCCTCTGCTACTCACTCTGAAAGCCAGACGAGGATGTGCTGGGGCCAG 3232  
QY 1053 GlySer-----PheProProGluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeu 1069  
DB 3233 GGCGCGCGCGGCTCTGCGCTCCGAGGCGCTGCGTGTGCGCACCAAGCATTCCTG 3292  
QY 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089  
DB 3293 CTCAGTGTGCTGACACCGCTGTCACCTAGTCCGCTCTCTGGGGTCACTCAGGACAGCC 3352  
QY 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109  
DB 3353 CAGACGAGCTGAGTCTGGAAGCTCCCGGGAGCAGCTGACTGCTGCGGCGCGCAGCC 3412  
QY 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122  
DB 3413 AACCGGCACTGCGCTCAGACTTCAAGACCATCCTGGAC 3451

## RESULT 12

US-10-044-539-1

Sequence 1, Application US/10044539

Publication No. US20030100093A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044,539

FILING DATE: 11-Jan-2002

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4015 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 56..3454

OTHER INFORMATION: /product= "hTRT"

/note= "human telomerase reverse

transcriptase (hTRT) catalytic protein

component"

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

-10-044-539-1

Alignment Scores:

Seq. No.:	0	Length:	4015
ore:	3505.00	Matches:	719
cent Similarity:	72.94%	Conservative:	122
st Local Similarity:	62.36%	Mismatches:	260
ery Match:	59.40%	Indels:	52
:	15	Gaps:	13

-09-042-460-2 (1-1122) x US-10-044-539-1 (1-4015)

```

1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
56 ATGCGCGCGCTCCCGCTCCGAGCGCGTCCCTGCTCGCGCAGCCACTACCGGAG 115
21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
116 GTGCTGCGGTGCGCAGTTCGTGCGCGCGCTGGGCGCCCGAGGCTGGCGGTGCGAG 175
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
176 CGCGGGAGCCCGCGGCTTTCGCGCGCTGGTGGCCAGTGTGCTGGTGGCGCTCGG 235
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLeuLeu 80
236 GAGCGACGGCGCGCCCGCGCGCTTCTCCGCGCAGGTGCTCTGCTGAGGAGGCTG 295
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
296 GTGCCCCGATGCTGCGAGAGCTGTGCGAGCGCGCGCGCGCGAGAGTGTGCGCTTCGCG 355
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
356 TTCGCGCTGTGAGCGGGCGCGCGGGCGCGCGGGCGCGCGGGCGCGCGGGCGCGGG 415
121 SerTyrLeuProAsnThrValIleGluThrArgValSerGlyAlaThrMetLeuLeu 140
416 AGCTACCTGCCCCAACCGGTGACCGAGCAGTCTCGGGGGAGCGGGGCGGTGGGCGTCTG 475

```

```

QY 141 LeuSerArgValGlyAspAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
DB 476 CTGCGCGCGGTGGCGGACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
QY 161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
DB 536 CTGCTGGCTCCAGCTCGCTACGAGTGTGGCGCGCGCTGTACGAGTGTGGCGCT 595
QY 181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
DB 596 GCCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 549
QY 201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerSerArgGlnGlnAlaPro 220
DB 650 -----TGCAGACGGCGCTGGAACCATAGCTCAGGAGGAGCGCGG 688
QY 221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
DB 689 GTCCCCCTGGCGCTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
QY 241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro--- 259
DB 749 CTGCGCTTGGCCAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 808
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
DB 809 GGGCAGGGCTCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGT 868
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerSerLys 292
DB 869 GTGGGTGCTCCTGCGAG-----CCGCCGAGAGAGCCACCTCTTTGGAG 913
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
DB 914 GGTGGCTCTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
DB 974 CCCCCATCCATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1033
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
DB 1034 GAGACCAAGCACTTCTCTACTCTCTCAGCGGAC---AAGGAGCAGCTGCGCGCGCTCT 1090
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIleIlePhe 370
DB 1091 CTACTCAGCTCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 1150
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
DB 1151 CTGGGTTCAGGCGCTGTGATCCGAGGAGTCCCGCGAGTTGCCCGCGCTGCCAGCGC 1210
QY 391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410
DB 1211 TACTGGCAATGCGCGCGCTGTTCTGAGAGTGTGTGGAAACACGCGCGAGTCCCGCTAC 1270
QY 411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla 430
DB 1271 GGGGTGCTCTCTCAGACGCACTGCGCGCTGCGAGCTGGG-----GTACCGCCAGCA 1321
QY 431 Leu-----Asn 432
DB 1322 GCCGGTGTCTGTGCGCGGAGAACGCCCGAGGGCTCTGTGGCGCGCGCGCGAGGAGGAG 1381
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
DB 1382 ACAGACCCCGGTGCGCTGTGTCAGCTGTCTCGCGCAGCAGCAGCAGCAGCAGCAGCTAC 1441
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
DB 1442 GGTCTGTCGCGCGCTGCTGCGCGGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1501

```



PRIOR FILING DATE: 2002-03-11

NUMBER OF SEQ ID NOS: 5

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1

LENGTH: 4027

TYPE: DNA

ORGANISM: ARTIFICIAL

FEATURE:

OTHER INFORMATION: Nucleotide sequence encoding telomerase reverse transcriptase

-10-385-882-1

Alignment Scores:

ad. No.: 0 Length: 4027

Matches: 3498.00

Conservative: 72.85%

Mismatch: 62.27%

Indels: 52

Gaps: 13

-09-042-460-2 (1-1122) x US-10-385-882-1 (1-4027)

```
1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
57 ATGCGCGCGCTCCCGCTGCGAGCGCTGCGCTCCCTGCTGCGCAGCCACTACCGGAG 116
21 ValTrpProLeuAlaThrPheValArgGlyProGluGlyArgGluValGln 40
117 GTGCTCCCTGCGCAGCTTCGTGCGCGCTGCGCGCCCGCCAGGCTGCGCTGTCGAG 176
41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
177 CGCGGGAGCCCGCGCTTTCGCGCGCTGCTGTCGCGCCAGTGCCTGGTGTGCTGCCCTGG 236
61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
237 GACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 296
81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100
297 GTGCGCGCGAGTGTGAGAGGTGTGCGAGCGCGCGCGCGAGACGCTGCTCGCTCGCG 356
101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
357 TTCGCGCTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 416
121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
417 AGTACTGCTCCCAACAGTGTACCGACGACCTGCGGGGAGCGCGCGCGCTGCGCGCTGCTG 476
141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160
477 TTGCGCGCGCTGCGCGACGAGTGTGTTACCTGCTGCGACGCTGCGCGCTGCTTTGTG 536
161 LeuValProProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180
537 CTGTGTGCTCCAGCTGCGCTACCGAGGTGTGCGCGCGCGCGCGCGCGCTGACGCTGCGCG 596
181 ThrThrAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200
597 GCGACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGA 650
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220
651 -----TCGCAACCGCGCGCTGGAACCATAGCTCAGGAGGCGCGCGG 689
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSerThrSer 240
690 GTCCCGCTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 749
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGluGlyPro 259
750 CTGCGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 809
```

```
QY 260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272
Db 810 GGGCAGCGGCTCTGCGCGCCACCGCGCAGGCGCGTGGACCGAGTGCAGTGGTTCTGT 869
QY 273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292
Db 870 GTGTGTACCTGCCAGA-----CCGCGCGAAGAGCCACCTCTTTGGAG 914
QY 293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysLysHisLysProSer 311
Db 915 GGTGCGCTCTCTGCGCAGCGCGCCACTCCACCCACTCGTGGCGCGCCACACCGCGGCG 974
QY 312 SerThrSerLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330
Db 975 CCCCATTCCATCGCGCGCCACCACTGCTCCCTGGGACACGCTTGTCCCCGGGTAGCGCC 1034
QY 331 GluThrArgHisPheLeuTyrSerArgGlyAspGlyGlnGluArgLeuAsnProSerPhe 350
Db 1035 GAGACCAAGCACCTTCTCTACTCTCTCAGCGCAGC---AAGGAGCAGCTGCGCGCTCTTC 1091
QY 351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgArgLeuValGluIlePhe 370
Db 1092 CTACTCAGCTCTCTGAGCGCCAGCCTGACTGGCGCTCGGAGGCTCTGGAGACCATCTTT 1151
QY 371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArg 390
Db 1152 CTGGTTCCAGCGCTCGATGTCAGGAGCTCCCGCAGGTTGCCCGCTGCCAGCGC 1211
QY 391 TyrTrpGlnMetArgProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyr 410
Db 1212 TACTGGCAATGCGCGCGCTGTTCTGAGAGCTGCTGGGAACACCGCAGGAGTCCCTAC 1271
QY 411 ValArgLeuLeuArgSerHisCysArgPheThrAlaAsnGlnGlnValThrAspAla 430
Db 1272 GGGTGTCTCTCAGACGCACTGCGCGCTGCGAGCTGCG-----GTCACCCCGAGCA 1322
QY 431 Leu-----Asn 432
Db 1323 GCGGTGTCTGTGCGCGCGGAGAGCCCGAGGCTCTGTGCGCGCGCGCGAGGAGGAG 1382
QY 433 ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452
Db 1383 ACAGACCCCGCTGCGCTGCTGCGAGCTGCTGCGCAGCAGCAGCCCTGGCAGGTTAC 1442
QY 453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472
Db 1443 GGTCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCGCTGCGCGCTGCGCGCTGCGCGC 1502
QY 473 AsnGluArgArgPhePheLysAsnLeuLysPheLysSerLeuGlyLysTyrGlyLys 492
Db 1503 AACGAACCGCGCTTCTCAGGAACACCAAGAGTTCTCTCTGGGAAGCATGCGCAG 1562
QY 493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512
Db 1563 CTCTCGCTCAGGAGCTGACGTGGAAGATGAGCTGCGGGCTGCGCTTGGCTGCGCAGG 1622
QY 513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluIleLeuAla 532
Db 1623 AGCCGAGGCTGCGCTGCTGCGCGCGCAGCAGCAGCTGCTGCGTGGAGAGATCCCTGG 1682
QY 533 ThrPheLeuPheTrpLeuMetAspThrTyrValGlnLeuLeuArgSerPhePheTyr 552
Db 1683 AGTTCCTGCTGCTGCTGATGAGTGTGATGCTGCTGAGCTGCTGCTCTTCTTTAT 1742
QY 553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572
Db 1743 GTCAGGAGACCATCTTCAAAGAACAGGCTCTTTTTCACCGAAGAGTGTCTGAGC 1802
QY 573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592
Db 1803 AGTTCGAAGAGATTTGGATCAGACGACCTTGAAGAGGTTGAGGCTGCGGAGCTGCG 1862
QY 593 GlnGluGluValArgHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612
```

```
1863 GAAGCAGAGGTCAGCAGCATCGGGAAGCCAGGCGCCCTGTGCTGCTCCAGCATCCGC 1922
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632
1923 TTCATCCCAAGCTTGAGCGGCTGCGCGCGATTGTGAACATGAGTACGTCGCGGGAGCC 1982
633 ArgAlaLeuGlyArgArgGlyGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652
1983 AGAACGTTCCGCGAGAGAAAGAGGCGCGAGCGTCTCACTCGAGGGTGAAGGCACTGTTT 2042
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672
2043 ACGGTGCTCAACTACGAGCGCGCGCGCGCGCGCTCTCTGCGCGCTCTGTGCTGCGG 2102
673 MetAsnAspIleTyrArgThrTyrArgAlaPheValLeuArgValArgAlaLeuAspGln 692
2103 CTGGAGCATATCCAGAGGCTTGGCGCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2162
693 ThrProArgMetTyrPheValAlaAspValThrGlyAlaTyrAspAlaIleProGln 712
2163 CGGCTGTAGCTGACTTGTGTCAAGGTGATGTGACGGCGCGGTACGACACCATCCCCAG 2222
713 GlyLysLeuValGluValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732
2223 GACAGGCTCACGAGGTATCGCCAGCATCATCAAA---CCCCAGAACAGTACTGCGTG 2279
733 ArgGlnTyrAlaValAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752
2280 COTCGGTATGCGGTGTCCAGAGGCGCGCCCATGGCGACGTCCGCAAGGCTTCAAGAGC 2339
753 GlnValThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772
2340 CACGTCTCTACCTTGACAGACTCCAGCGGTACATGCGACAGTTCGTGTGCTCACTGCG 2399
773 AspSerAspAlaSerAlaLeuArgAsnSerValIleGluGlnSerIleSerMetAsn 792
2400 GAGACC-----AGCCCGCTGAGGATGCCCTGCTCATCGAGCAGAGCTCTCTCTGAT 2453
793 GluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerValValLys 812
2454 GAGGCGCAGGTGGCTCTTGGAGCTCTTCTGAGCTTCTGAGCTTCTGAGCTTCTGAGCT 2513
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySerSerLeuSerThr 832
2514 ATCAGGGCGAGCTCTGCTGCTGAGTGGCGGAGTCCGCGAGGCTCCATCTCTCTCAGC 2573
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852
2574 CTGCTCTGAGCTGTGCTACGGCGACATGAGAACAAAGCTGTTTTCGGGGATTCGGCGG 2633
853 AspGlyLeuLeuLeuArgPheValAspPheLeuValThrProHisLeuAspGln 872
2634 GACGGGCTGCTCTGCTGCTGAGTATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2693
873 AlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCysMetIleAsn 892
2694 GCGAAACCTTCTCTCAGACCTGTGTCGAGGTGTCTCTGAGTATGCTGCGGTGTGAGC 2753
893 LeuGlnLysThrValValAsnProValGluProGlyThrLeuGlyGlyAlaAlaPro 912
2754 TTGCGGAAGACAGTGTGTGAATTCCTCTAGAACAGGAGGCGCTGCGGTGCGGCTTTT 2813
913 TyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThr 932
2814 GTTCAGATCGCGGCGCGCGCTATTCCTCTGCTGCGGCTGTGCTGTGATACCCGAGC 2873
933 LeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThrSerLeuThr 952
2874 CTGGAGGTGCAGAGCGACTACTCTCAGCTATGCCGAGCTTCCATCAGAGCGAGTCTCAC 2933
953 PheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSerValLeuArg 972
```

```
Db 2934 TTCAACCGCGCTTCAAGGCTGAGGAGAAATCGTGTGCAAACTCTTTGGGGTCTTGGCG 2993
Qy 973 LeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIle 992
Db 2994 CTGAAGTGTCAAGCCCTGTTTCTGGATTTCAGGTGACAGCTTCCAGACGGTGTGACCC 3053
Qy 993 AsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeu 1012
Db 3054 AACATCTACAAGATCTCTCTGCTGCTGAGGCGTACAGGTTTCAAGCATGTGTGTGACGCTC 3113
Qy 1013 ProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGln 1032
Db 3114 CCATTTCATCAGCAAGTTTGGAAAGACCCACATTTTCTGCGCGCTCATCTCTGACAGC 3173
Qy 1033 AlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeuLysAlaSer 1052
Db 3174 GCCTCCCTCTGCTACTCTCATCTCTGAAAGCCAGAGGATGTGCTGGGGGCAAG 3233
Qy 1053 GlySer-----PheProGluAlaAlaHisTyrPleuCysTyrGlnAlaPheLeu 1069
Db 3234 GCGCCCGCGGCTCTGCGCTCCGAGCGCGTGCAGTGGCTGTGCCACCAAGCATTCCTG 3293
Qy 1070 LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAla 1089
Db 3294 CTCAGCTGTACTGCACACCGTGTCACTAGTCCCACTCTCTGGGGTCACTCAGGACAGC 3353
Qy 1090 GlnLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109
Db 3354 CAGACGAGCTGAGTCGGAAGCTCCCGGGGACGACGCTGACTGCCCTGGAGCGCGAGCC 3413
Qy 1110 AspProAlaLeuSerThrAspPheGlnThrIleLeuAsp 1122
Db 3414 AACCGGCACTGCGCTCAGACTTCAAGACCATCTCTGGAC 3452
```

## RESULT 14

```
US-10-205-629-1
; Sequence 1, Application US/10205629
; Publication No. US20030049236A1
; GENERAL INFORMATION:
; APPLICANT: Kassem, Moustapha
; APPLICANT: Jensen, Thomas
; APPLICANT: Rattan, Suresh
; TITLE OF INVENTION: Immortalized Stem Cells
; FILE REFERENCE: 006148.00002
; CURRENT APPLICATION NUMBER: US/10/205,629
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/315939
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: PA 2001 01148
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-205-629-1

Alignment Scores:
Pred. No.: 0 Length: 3453
Score: 3496.00 Matches: 717
Percent Similarity: 72.85% Conservative: 123
Best Local Similarity: 62.19% Mismatches: 261
Query Match: 59.24% Indels: 52
DB: 15 Gaps: 13

US-09-042-460-2 (1-1122) x US-10-205-629-1 (1-3453)
Qy 1 MetThrArgAlaProArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20
Db 10 ATGCGCGGCGCTCCCGCTGCGGAGCGGTGCGCTCTCTGTCGCGCAGCCACTACCGCGAG 69
Qy 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgArgLeuValGln 40
```

Db	1105	CTGGGTTCCAGGCCCTCGATGTCAGGAGACTCCCGCAGGTGTGCCCGCTGCCCCGCTGCCCCAGCGC	1161
Qy	391	TyTTPGlnMetArgProLeuPheGlnGlnLeuLeuValAsnHisAlaGluCysGlnTyr	410
Db	1165	TACTGGCAATATGCGGCCCTGTTCCTGGAGAGTGTCTGGGAACCAACGCGCAGTGCCTCCATC	1224
Qy	411	ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAla	430
Db	1225	GGGGTGTCTCTCAAGACGCACTGCCCGCTGCAGCTGCG-----GTCAACCCAGCA	1275
Qy	431	Leu-----	432
Db	1276	GCCGGTGTCTGTGCCCGGAGAACCCAGGCGCTGTGTGGCGGCCCGCCAGGAGGAGAC	1335
Qy	433	ThrSerProProHisLeuMetAspLeuLeuArgLeuHisSerSerProTTPGlnValTyr	452
Db	1336	ACAGACCCCGCTGCCCTGGTGCAGCTGTCCCGCAGCAGCAGCAGGCCCTCTGCAGGTGTAC	1395
Qy	453	GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTTPGlyThrArgHis	472
Db	1396	GGCTTCGTGGGGCTCTGCTGCGCGGCTGTGCCCCAGGCCCTCTGGGGCTCCAGGCAC	1455
Qy	473	AsnGluArgArgPhePheLysAsnLeuLysLysPheLysSerLeuGlyIleTyrGlyLys	492
Db	1456	AACGAACCGCGCTTCTCAGAAACCAAGAGTTTCATCTCCCTGGGGAAGCATGCCAAG	1515
Qy	493	LeuSerLeuGlnGluLeuMetTTPLysMetLysValGluAspCysHisTTPLeuArgSer	512
Db	1516	CTCTCGCTCGAGGAGCTGACGTGGAGATGAGCGTGGCGGACTGGCGCTTGGCTCGGCAGG	1575
Qy	513	SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlyLeuAla	532
Db	1576	AGCCAGGGGTGGCTGTGTTCGGCGCAGCAGACACGCTCTCGTGAGAGATCCTCGCC	1635
Qy	533	ThrPheLeuPheTTPLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr	552
Db	1636	AAGTTCCTGCACTGCCTGATGAGTGTACGTCTGTCGAGCTGCTCAGTCTTCTTTAT	1695
Qy	553	IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTTPSer	572
Db	1696	GTCCGGAGACCAGTTTCAAAAGAACAGGCTCTTTTCTACGGCCGAGTGTCTGGAGC	1755
Qy	573	LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer	592
Db	1756	AAGTTCGAAAGCATTGGAACTCAGACAGCACTTGAAGAGGGTGCAGCTCGGAGAGCTGCG	1815
Qy	593	GlnGluGluValArgHisHisGlnAspThrTTPLeuAlaMetProIleCysArgLeuArg	612
Db	1816	GAACAGAGGTTCAGCAGCATCCGGAAGCCAGCCCGCCCTCTGCTACGCTCCAGACTCCGC	1875
Qy	613	PheIleProLysProAenglyLeuArgProIleValAsnMetSerTyrSerMetGlyThr	632
Db	1876	TTCATCCCAAGCCTGACGGGCTGCGGCGAATTGTGAACATGCATGCTCGTGGGAGCC	1935
Qy	633	ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLysThrLeuPhe	652
Db	1936	AGACGTCCCGAGAAAGAGGCGCAGCGTCTCACCCTCAGAGGGTGAAGGCATGTTTC	1995
Qy	653	SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly	672
Db	1996	AGCGTGTCAACTCAGAGCGGCGCGCGGCCCGCCGCTCTCTGGGGCGCTCTGTGTGGGC	2055
Qy	673	MetAsnAspIleTyrArgThrTTPArgAlaPheValLeuArgValArgAlaLeuAspGln	692
Db	2056	CTGGACGATATCCACAGGCCCTGGCGCACCTCTGCTGCTGCTGCGGCGCCAGGACCCG	2115
Qy	693	ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln	712
Db	2116	CGCGCTGACGTACTTTGTCAAGGTGTGATGTACGGCGCGCTGACGACCATCCCCAG	2175
Qy	713	GlyLysLeuValGluValValAlaAsnMetIleArgHisSerGluSerThrTyrCysIle	732
Db	2176	GACAGGCTCAGGAGGTCACTCCGACATCATCAAA---CCCCAGAACCACTACTGGGT	2232





141 LeuSerArgValGlyAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
4409 CTGCGCGCGTGGCGACGAGCTGTGGTTCACTCTGTGGCAGCTCGCGCTCTTTGTG 4468  
161 LeuValProSerCysAlaTyrGlnValCysGlySerProLeuTyrGlnIleCysAla 180  
4469 CTGGTGGTCTCCAGCTGGCTTACAGGTGTGGCGCGCGCTGTACCAAGCTCGCGCT 4528  
181 ThrTrpAspIleTrpProSerValSerAlaSerTyrArgProThrArgProValGlyArg 200  
4529 GCCACTCAGCGCGCGCGCGCACACGCTAGT---GGACCCCGAAGCGCTCTGGGA--- 4582  
201 AsnPheThrAsnLeuArgPheLeuGlnGlnIleLysSerSerArgGlnGluAlaPro 220  
4583 -----TGCAACGGCGCTGGAAACCATACGTCGAGGAGCGCGG 4621  
221 LysProLeuAlaLeuProSerArgGlyThrLysArgHisLeuSerLeuThrSer 240  
4622 GTCCCTCTGGCGCTGCCAGCCCGGTCGAGGAGCGCGCGGCGAGTGCACCGCAAGT 4681  
241 ValProSerAlaLysLysAlaArgCysTyrProValProArgValGluGlyPro--- 259  
4682 CTGCGGTCTCCCAAGAGCGCCAGCGTGGCGTGCCTGAGCGCGAGCGCGCGCTT 4741  
260 -----HisArgGlnValLeuProThrProSerGlyLysSerTrp--- 272  
4742 GGGCAGGGTCTGGCGCCACCGGGGAGGAGCGGTGACCGAGTACCGTGTCTTCTGT 4801  
273 ValProSerProAlaArgSerProGluValProThrAlaGluLysAspLeuSerLys 292  
4802 GTGGTGTACCTGCCAGA-----CCGCGCGAAGACCACTCTTTGGAG 4846  
293 GlyLysValSerAspLeuSerLeuSer---GlySerValCysCysAlaHisLysProSer 311  
4847 GGTGCGCTCTGTGGCAGCGCGCCACTCCACCCCTCCGTGGCGCCGACACCGCGGGC 4906  
312 SerThrLeuLeuSerProProArgGlnAsnAlaPheGlnLeuArgPro---PheIle 330  
4907 CCCCACCATCGCGCGCCACACGCTCCCTGGGACACGCTTGTCCCGCGGTACGCC 4966  
331 GluThrArgHisPheLeuTyrSerArgLysAspGlyGlnGluArgLeuAsnProSerPhe 350  
4967 GAGACAAGACCTTCTCTACTCTCTCAGCGAC---AAGGAGCAGCTCGCGCCCTCTTC 5023  
351 LeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArgLeuValGluIlePhe 370  
5024 CTACTCAGCTCTGTAGGCCAGCGCTGACTGGCGCTCGAGGCTCGTGGAGACCATCTT 5083  
371 LeuGlySerArgProArgThrSerGlyProLeuCysArgThrHisArgLeuSerArg 390  
5084 CTGGGTTCAGGCCCTGGATGCCAGGACTCCCGCAGGTTGCCCGCGCTCCCGCGCGC 5143  
391 TyrTrpGlnMetArgProLeuPheGlnLeuLeuValAsnHisAlaGluCysGlnTyr 410  
5144 TACTGCAAAATGGGCGCTCTCTGAGCTGCTTGGAGACCGAGCGAGTGGCCCTAC 5203  
411 ValArgLeuLeuArgSerHisCysArgPheArgThrAlaAsnGlnValThrAspAla 430  
5204 GGGGTGTCTCTCAAGACGCACTGCCGCTGCGAGTGGG-----GTCACCCCGACA 5254  
431 Leu-----Asn 432  
5255 GCGGTGTCTGTGCGGAGAGAGCCCGAGGCTCTGTGGCGCGCCCGAGGAGGAC 5314  
433 ThrSerProHisLeuMetAspLeuLeuArgLeuHisSerSerProTrpGlnValTyr 452  
5315 ACAGACCCCGCTGCTGTGAGCTGTCTCCGCGACACAGCAGCGCCCTGGCAGGTGAC 5374  
453 GlyPheLeuArgAlaCysLeuCysLysValValSerAlaSerLeuTrpGlyThrArgHis 472  
5375 GGCTTGTGGCGGCTGTGCGCGCGCTGGTGGTGGCGCGCGCGCGCGCGCGCGCGCG 5434

473 AsnGluArgArgPhePheLysAsnLeuLysLysPheIleSerLeuGlyLysTyrGlyLys 492  
5435 AACGAACCGCGCTTCTCAGGAACACCAAGAGTTTCATCTCCTGGGAGCATGCCAAG 5494  
493 LeuSerLeuGlnGluLeuMetTrpLysMetLysValGluAspCysHisTrpLeuArgSer 512  
5495 CTCTCGCTGAGGAGCTGAGTGGAGATGAGCGTGGGACTGCGCTGGCTGGCGCAGG 5554  
513 SerProGlyLysAspArgValProAlaAlaGluHisArgLeuArgGluArgIleLeuAla 532  
5555 AGCCACGGGTGTGGTGTGTCCGGCGCAGACACCGCTCTCGTGAGGAGATCCTGGCC 5614  
533 ThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePheTyr 552  
5615 AAGTCTCTGACCTGGCTGAGTGTAGCTGTAGCTGTAGCTGTCTCAGTCTTTCTTTAT 5674  
553 IleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrpSer 572  
5675 GTACGGAGACCACTTCAAAAGAACAGGCTCTTTTCTACGGCGGAGTGTCTGGAGC 5734  
573 LysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGluLeuSer 592  
5735 AAGTTGCAAGCAATGGATCAGACAGCACTTGAAGAGGTGAGCTGGGAGCTGTG 5794  
593 GlnGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeuArg 612  
5795 GAAGCAGAGGTTCAGGACGATCGGAAGACGAGCGCGCGCTGTGACGCTCCAGACTCCGC 5854  
613 PheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGlyThr 632  
5855 TTATCCCGAAGCTGACGGCTGCGCGCGATTTGAATGATGATGATGATGATGATGATG 5914  
633 ArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeuPhe 652  
5915 AGAAGCTCCCGACAGAAAGAGCGCGCGCTCTCACCTCGAGGCTGAAGCACTGTTC 5974  
653 SerMetLeuAsnTyrGluArgThrLysHisProHisLeuMetGlySerSerValLeuGly 672  
5975 AGCGTGTCAACTACGAGCGCGCGCGCGCGCTCTCGGGCGCTCTGTGTGGGCG 6034  
673 MetAsnAspIleTyrArgThrTrpAlaPheValLeuArgValArgAlaLeuAspGln 692  
6035 CTGACGATATCCACAGGCTGGCGCACTTCTGTCTGCTGTGGCGCCGAGGACCG 6094  
693 ThrProArgMetTyrPheValLysAlaAspValThrGlyAlaTyrAspAlaIleProGln 712  
6095 CCGCTGAGCTGTACTTGTCAAGGTGATGTGACGGCGCTGACGACCACTCCCGCAG 6154  
713 GlyLysLeuValGluValAlaAlaAsnMetIleArgHisSerGluSerThrTyrCysIle 732  
6155 GACAGGCTCAGGAGGTTCATCGCCAGCATCATCAA---CCCCAGAACAGTACTCGTG 6211  
733 ArgGlnTyrAlaValArgAspSerGlnGlyGlnValHisLysSerPheArgArg 752  
6212 CGTGGTATCGGTGTTCAGAGCGCGCCATCGGCGACGTCGCGCAAGCGCTTCAAGAGC 6271  
753 GlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuLysHisLeuGln 772  
6272 CACGCTCTACTCTTGACAGACCTCCAGCGGTACATGCGACAGTCTGGGCTCCTCCAG 6331  
773 AspSerAspAlaSerAlaLeuArgAsnSerValValIleGluGlnSerIleSerMetAsn 792  
6332 GAGACC-----AGCCGCTGAGGATCGCTCATCGAGCAGAGCTCCTCCCTGAAT 6385  
793 GluSerSerSerLeuPheAspPheLeuHisPheLeuArgHisSerValValLys 812  
6386 GAGGACAGCAGTGGCTCTTCAGCTTCTTCTACGCTTCATGTGCCACCAAGCGCGTGGC 6445  
813 IleGlyAspArgCysTyrThrGlnCysGlnGlyProGlnGlySerSerLeuSerThr 832  
6446 ATCAGGCGCAGTCTCTAGCTGCGAGGGATCCCGCAGGGCTCCATCTCTCCAGC 6505  
833 LeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAlaGluValGlnArg 852



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - nucleic search, using frame\_plus\_p2n model

on: February 1, 2004, 15:05:59 ; Search time 5256 Seconds  
(without alignments)  
5188.284 Million cell updates/sec

le: US-09-042-460-2  
fect score: 5901  
uence: 1 NTRAPCPAVALSLRLSYRE.....TILKAADPALSTDFOTILD 1122

ring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

urched: 22781392 seqs, 12152238056 residues

al number of hits satisfying chosen parameters: 45562784

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

mand line parameters:

>DEL=frame\_p2n.model DEV=xlh  
>DEL=1/USPTO\_spool/US09042460/runat\_01022004\_150551\_17026/app\_query.fasta\_1.1297  
i=EST -QFMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
ITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
>ALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
>FMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
>SER=US09042460 @CGN 1.1 3909 @runat\_01022004\_150551\_17026 -NCPU=6 -ICPU=3  
>WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
>IV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
>APEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEX=7

abase :  
1: em estba.\*  
2: em esthum.\*  
3: em estin.\*  
4: em estmu.\*  
5: em estov.\*  
6: em estpl.\*  
7: em estro.\*  
8: em htc.\*  
9: gb estl.\*  
10: gb est2.\*  
11: gb htc.\*  
12: gb est3.\*  
13: gb est4.\*  
14: gb est5.\*  
15: em estfun.\*  
16: em estom.\*  
17: em ges hum.\*  
18: em ges inv.\*  
19: em ges pln.\*  
20: em ges vrt.\*  
21: em ges fun.\*  
22: em ges mam.\*  
23: em ges mus.\*  
24: em ges pro.\*  
25: em ges rod.\*  
26: em ges phg.\*  
27: em ges vrl.\*  
28: gb gssl.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1455	24.7	851	13	BU702370	BU702370 UI-M-F10-
2	1279.5	21.7	851	12	BG917907	BG917907 602820830
3	1173	19.9	664	13	BQ258274	BQ258274 NISC kp11
4	1082	18.3	641	28	AZ972318	AZ972318 2M0246F07
5	938.5	15.9	925	12	BM453198	BM453198 AGENCOURT
6	841	14.3	599	10	BB618671	BB618671 BB618671
7	783	13.3	866	10	BB371943	BB371943 601217728
8	768.5	13.0	614	10	BB651920	BB651920 BB651920
9	595	10.1	347	9	AW244516	AW244516 BR END06B6
10	558	9.5	492	12	BM824748	BM824748 K-EST0096
11	540.5	9.2	753	13	BU452535	BU452535 603767927
12	511	8.7	326	13	BY149368	BY149368 BY149368
13	493.5	8.4	775	12	BT388013	BT388013 BFL26_002
14	483	8.2	880	13	BU377259	BU377259 603811228
15	463	7.8	389	9	AA281296	AA281296 zt08902.r
16	462	7.8	668	14	CA380121	CA380121 659344 NC
17	435	7.4	696	13	BU139751	BU139751 603134527
18	427.5	7.2	632	14	CA353864	CA353864 625469 NC
19	407	6.9	568	4	BX521269	BX521269 RZPD Mus
20	405	6.9	534	9	AW318894	AW318894 un09a02.y
21	396.5	6.7	835	13	BU111946	BU111946 603127372
22	392.5	6.7	875	13	BU122597	BU122597 603148441
23	337	5.7	715	10	BE396925	BE396925 601290610
24	335	5.7	610	10	BE514188	BE514188 601316376
25	328	5.6	649	10	BE514070	BE514070 601316575
26	327	5.5	409	9	AA311750	AA311750 EST182469
27	325.5	5.5	813	13	BU224024	BU224024 603798349
28	321	5.4	679	10	BE396606	BE396606 601289077
29	293	5.0	779	10	BE268183	BE268183 601125261
30	282.5	4.8	1455	29	CC190951	CC190951 CH261-380
31	279.5	4.7	1230	29	CC211572	CC211572 CH261-14F
32	279.5	4.7	1554	29	CC190875	CC190875 CH261-38M
33	272	4.6	554	28	AQ397020	AQ397020 m9kb0014C
34	266	4.5	1020	29	CNS05F1M	AL334435 Tetraodon
35	250.5	4.2	570	12	BM521744	BM521744 sak70f07
36	244.5	4.1	502	10	BE251764	BE251764 EST419026
37	216.5	3.7	487	28	AZ799615	AZ799615 2M0057L07
38	214.5	3.6	739	13	BK315053	BK315053 BX315053
39	212	3.6	696	28	BH962910	BH962910 odj2ib08
40	209.5	3.6	393	13	BK088059	BK088059 BX088059
41	205.5	3.5	393	10	BF251775	BF251775 EST419037
42	203.5	3.4	625	28	B27802	B27802 T1704TRD TA
43	198	3.4	613	10	BF597086	BF597086 sub3903.Y
44	197	3.3	1009	29	CNS0730U	AL427780 clone BA0
45	167.5	2.8	724	13	BK315052	BK315052 BX315052

ALIGNMENTS

RESULT 1  
BU702370  
LOCUS BU702370 851 bp mRNA linear EST 09-OCT-2002  
DEFINITION UI-M-F10-byx-f-12-0-UI.r1 NIH\_BMAP\_F10 Mus musculus cDNA clone  
IMAGE: 6400523 5', mRNA sequence.  
ACCESSION BU702370  
VERSION BU702370.1 GI:23627105  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 851)

# AUTHORS TITLE JOURNAL COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
This clone was contributed by the Brain Molecular Anatomy Project  
(BMAP)

Seq primer: pYX-5.

## FEATURES

### source

Location/Qualifiers  
1. 851  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CS7BL/6"  
/db\_xref="taxon:10090"  
/clone="IMAGE: 6400523"  
/tissue="whole brain"  
/dev\_stage="embryo 12.5dpc"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_lib="NIH\_BMAP\_F10"  
/note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with EcoR I adaptor, digested with NotI and then cloned  
directionally into pYX-Asc vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CAGCCACGAC. This library was created for the University  
Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
Developing Mouse Nervous System', supported by National  
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
program coordinator."

SE COUNT 180 a 227 c 221 g 220 t 3 others

### IGN

### ignment Scores:

ed No.: 1.17e-118 Length: 851  
ore: 1455.00 Matches: 278  
cent Similarity: 98.58% Conservative: 0  
st Local Similarity: 98.58% Mismatches: 4  
ery Match: 24.66% Indels: 0  
Gaps: 0

-09-042-460-2 (1-1122) x BU702370 (1-851)

729 ThrTyrCysIleArgGlnTyrAlaValAlaValArgAspSerGlnGlyGlnValHisIys 748  
6 ACCTACTGTATCCGACAGTATGATGAGTGGTCCGAGAGATAGCCAGGCCAAGTCCACAG 65  
749 SerPheArgGlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeu 768  
66 TCCTTTAGGAGACAGTCCACCCCTCTGACCTCCAGCCATACATGGCCAGTTCCT 125  
769 LysHisLeuGlnAspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSer 789  
126 AAGCATCTGCAGGATTTCAGATGCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 185  
789 IleSerMetAsnGlnSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHis 808  
186 ATCTCTAAGTATGAGACACACACCCCTTTGACTTCTTCTGCACTTCTGCTGC 245  
809 SerValValIleGlyAspArgCysTyrThrGlnCysGlnGlyIleProGlnGlySer 828  
246 AGTGTCTGAAGATTGGTGAAGTGTATACGAGTGCAGGCGATCCCGGCGCTCC 305

QY 829 SerLeuSerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLysLeuPheAla 848  
Db 306 AGCCTATCCACCTGCTCTGCACTGTGTTTCGGAGACATGGAGACAAGCTGTTGCT 365  
QY 849 GluValGlnThrArgAspGlyLeuLeuLeuArgPheValAspAspPheLeuValThrPro 868  
Db 366 GAGTGTGAGCGGATGGGTGCTTTTACGTTTGTGAGTCTTTCTGTTGTTGAGCTCT 425  
QY 869 HisLeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGly 888  
Db 426 CACTTGGACCAAGCAAAAACCTTCTCCAGCACCTGCTCATGGGCTTCTCAGTATGG 485  
QY 889 CysMetIleAsnLeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGly 908  
Db 486 TGCATGATAAAGTTGCAGAGACAGTGGTGAACCTCCCTGTGAGCTGTACCTGGT 545  
QY 909 GlyAlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeu 928  
Db 546 GTGCAGCTCCATACAGCTGCTGCTCACTGCTGTTCCCTGGTGTGGCTTGTGCTGTG 605  
QY 929 AspThrGlnThrLeuGlnValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLys 948  
Db 606 GACACTCAGACTTTGGAGGTGTTCTGTGACTACTCAGGTATGCCAGACCTCAATTAG 665  
QY 949 ThrSerLeuThrPheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeu 968  
Db 666 ACAGGCTCCACTCCAGAGTCTTCANAGCTGGGAGACCATGGGAAACANGCTCCTG 725  
QY 969 SerValLeuArgLeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGln 988  
Db 726 TCGGCTCTGGGTTGAAGTGTCAAGTGTCAAGTGTATTTCTAGACTTCGAGGTGAACAGCTCCAG 785  
QY 989 ThrValCysIleLeuIleTyrLysIlePheLeuGlnAlaTyrArgPheHisAlaCys 1008  
Db 786 ACAGTGTGCATCAATATATACAGATCTTCTGCTTCANGGCTACAGGTTCATGTCATGT 845  
QY 1009 Vallile 1010  
Db 846 GTGATT 851

### RESULT 2

### BG917907

### LOCUS

### DEFINITION

### ACCESSION

### VERSION

### KEYWORDS

### SOURCE

### ORGANISM

### REFERENCE

### AUTHORS

### TITLE

### JOURNAL

### COMMENT

### FEATURES

### source

### 1. 851

### /organism="Mus musculus"

### /mol\_type="mRNA"

### /strain="FVB/N"

### 602820830F1 NCI\_CGAP\_Mam6 Mus musculus cDNA clone IMAGE:4949887 5',

### mRNA sequence.

### EST.

### Mus musculus

### Mus musculus

### Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

### Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

### 1 (bases 1 to 851)

### NIH-MGC <http://mgc.nci.nih.gov/>.

### National Institutes of Health, Mammalian Gene Collection (MGC)

### Unpublished

### Contact: Robert Strausberg, Ph.D.

### Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

### Tissue Procurement: Jeffrey Green M.D.

### cDNA Library Preparation: Life Technologies, Inc.

### cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

### DNA Sequencing by: Incyte Genomics, Inc.

### Clone distribution: MGC clone distribution information can be

### found through the I.M.A.G.E. Consortium/LLNL at:

### <http://image.llnl.gov>

### Plate: LLAM10903 row: k column: 08

### High quality sequence stop: 753.

### Location/Qualifiers

/db xref="taxon:10090"  
 /clone="IMAGE:4949887"  
 /sex="female, virgin"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /dev\_stage="5 months"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Mam6"  
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;  
 Site 2: NotI; Cloned unidirectionally; Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigator  
 providing samples: Jeffrey Green, M.D., NIH"  
 E COUNT 184 a 238 c 214 g 215 t  
 GIN

gment Scores:  
 d. No.: 4, 71e-103 Length: 851  
 re: 1279.50 Matches: 258  
 cent Similarity: 95.93% Conservative: 1  
 t Local Similarity: 95.56% Mismatches: 10  
 ry Match: 21.68% Indels: 4  
 12 Gaps: 0

09-042-460-2 (1-1122) x BG917907 (1-851)

850 ValGlnArgAspGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuValThrProHis 869  
 13 GTGCAGCGGGATGGGTG-CATTTACGTG-TGTGATGACTTTCTGTAGTGACGGCTCAC 70  
 870 LeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCys 889  
 71 TTGACCAAGCAAAACCTTCTCCAGCACCTGGTCTCATGGCGTTCTGAGTATGGGTGC 130  
 890 MetIleAsnLeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGly 909  
 131 ATGATAAATTGCAGAGACAGTGGTGAATCTCCCTGTGGAGCTGGTACCTGGGTGCT 190  
 910 AlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuAsp 929  
 191 GCAGTCCATACCAAGCTGCTGCTCACTGCTGTTTCCCTGGTGGCTGTGCTGGAC 250  
 930 ThrGlnThrLeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLysThr 949  
 251 ACTCAGACTCTGGAGGTGTTCTGTGACTACTCAGGTATATGCCAGACCTCAATTAGAGC 310  
 950 SerLeuThrPheGlnSerValPheLysAlaGlyLysThrMetArgAsnLysLeuLeuSer 969  
 311 AGCTCACTTCCAGAGTGTCTTCAAAGCTGGGAAGACCATGCGGAACAAGCTCCTGTGCG 370  
 970 ValLeuArgLeuLysCysHisGlyLeuPheLeuAspLeuGlnValAsnSerLeuGlnThr 989  
 371 GTCTTGGGTGAAGTGTACGGTCTATTTCTAGACTTGCAGGTGAACAGCTCCAGACA 430  
 990 ValCysIleAsnIleTyrLysIlePheLeuLeuGlnAlaTyrArgPheHisAlaCysVal 1009  
 431 GTCTGCATCAATATATCAAGATCTTCTCCTCTCAGGCTTACAGGTTCATGATGATGTG 490  
 1010 IleGlnLeuProPheAspGlnArgValArgLysAsnLeuThrPhePheLeuGlyIleIle 1029  
 491 ATTCAAGCTTCTCTTGGACCGGTGTAGAAAGACCTCACATCTTCTTGGGCATCATC 550  
 1030 SerSerGlnAlaSerCysCysTyrAlaIleLeuLysValLysAsnProGlyMetThrLeu 1049  
 551 TCCAGCCAAGCATCTGCTGTATGCTATCTCTCAAGGTCAAGATTCAGGAATGACACTA 610  
 1050 LysAlaSerGlySerPheProGluAlaAlaHisIleTyrLeuCysTyrGlnAlaPheLeu 1069  
 611 AAGGCCCTGTGCTCTCTCTCTCTGAGCGCA-CATTTGGCTGTGCTACAGGCCCTTCCTG 669  
 1070 -LeuLysLeuAlaAlaHisSerValIleTyrLysCysLeuLeuGlyProLeuArgThrAl 1089  
 670 GCTCAAGCTGGTCTCTATCTGTCTCATCTCAAAATGTCTCTGGGACCTCTGAGGACAGC 729  
 1089 agLysLeuLeuCysArgLysLeuProGluAlaThrMetThrIleLeuLysAlaAla 1109

Db 730 CCAAAACCGCTGTCCGGAAGCTCCAGAGCGGACATGACCATGCTTACAGTGCAGC 789  
 Qy 1109 aAspProAlaLeuSerThrAspPheGln 1118  
 Db 790 TGACCCAGACCTAAGGACAGACTTTTCAG 817

RESULT 3  
 BQ258274 664 bp mRNA linear EST 06-MAY-2002  
 LOCUS NISC kp1lg04.q3 Baker mouse embryo e7.5 Mus musculus cDNA clone  
 BQ258274 IMAGE:5409222, mRNA sequence.  
 ACCESSION BQ258274  
 VERSION BQ258274.1 GI:20459030  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 REFERENCE 1 (bases 1 to 664)  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs@mail.nih.gov  
 cDNA Library Preparation: J. Baker (Stanford University)  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 info@image.llnl.gov  
 MGI:1845958  
 Plate: LLAM12043 row: N column: 7  
 Seq primer: Sp6 primer.  
 Location/Qualifiers  
 1..664  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD-1"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5409222"  
 /tissue\_type="embryo, late gastrula"  
 /dev\_stage="embryo, 7.5 dpc"  
 /lab\_host="XLI-Blue"  
 /note="Vector: pCG105; Site 1: NotI; Site 2: Sali; cDNA  
 made by oligo-dT priming. Directionally cloned into  
 Sali/NotI sites using the following 5' adaptor:  
 5'-TCGACCCAGCGCTCG-3'. Size-selected for average insert  
 size 1.8-1.9 kb. Library constructed by J. Baker (Stanford  
 University)."  
 BASE COUNT 155 a 177 c 183 g 149 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 9,31e-94 Length: 664  
 Score: 1173.00 Matches: 218  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.54% Mismatches: 0  
 Query Match: 19.88% Indels: 0  
 DB: 13 Gaps: 0

US-09-042-460-2 (1-1122) x BQ258274 (1-664)

Qy 356 GlnProAsnLeuThrGlyAlaArgLeuValGluIlePheLeuGlySerArgPro 375  
 Db 7 GAGCCTAACTTGCATGGGCGCCAGGAGCTGGTGGAGATCATCTTCTGGGCTCAAGGCT 66  
 Qy 376 ArgThrSerGlyProLeuCysArgThrHisArgLeuSerArgArgTyrTrpGlnMetArg 395  
 Db 67 AGGACATCAGGACCACTCTGCGAGGACACACCGCTCTATCGCGTCGATACTGCGCAGATCGG 126

```

336 ProLeuPheGlnGlnLeuValAsnHisAlaGluCysGlnTyrValArgLeuLeuArg 415
127 CCCCTGTTCCACAGCTCTCGTGAACCATCAGAGTCCCAATATGTCAGACTCCCTCAG 186
416 SerHisCysArgPheArgThrAlaAsnGlnGlnValThrAspAlaLeuAsnThrSerPro 435
187 TCACATTGCAGGTTTCGACAGCAGCAACCAACAGGTGACAGTSCCTTGACACAGCCCA 246
436 ProHisLeuMetAspLeuLeuArgLeuHisSerSerProTyrPheGlnValTyrGlyPheLeu 455
247 CGCACCTCATGATTGCTCGCTGACACAGAGTCCCTGGCAGGTATATGTTTCTT 306
456 ArgAlaCysLeuCysValValSerAlaSerLeuTyrGlyThrArgHisAsnGluArg 475
307 CGGGCTGCTCTCGAAGGTGTCTGCTAGTCTCTGGGGTACCAGGACATGAGCGC 366
476 ArgPhePheLeuAsnLeuLysPheLeuSerLeuGlyLysTyrGlyLysLeuSerLeu 495
367 CGCTTCTTTAAGAACTTAAGAAAGTTTCATCTCGTTGGGAAATACGGCAAGCTATCACTG 426
496 GlnGluLeuMetTyrLysMetLysValGluAspCysHisTyrPheLeuArgSerProGly 515
427 CAGGAACTGATGTGAAGATGAAGATGAGGATTGCCACTGGCTCCGACAGGCCAGGG 486
516 LysAspArgValProAlaAlaGluHisArgLeuArgGluArgGlnLeuAlaThrPheLeu 535
487 AAGGACCGTGTCCCGCTCCAGAGCACCGTCTGAGGGAGAGGATCTGCTAGTTCCTG 546
536 PheTyrLeuMetAspThrTyrValValGlnLeuLeuArgSerPheTyrIleThrGlu 555
547 TTCTGCTGATGGACACATACGTGTGATAGCTGCTTAGTTCATTTTATCATCACAGAG 606
556 SerThrPheGlnLysAsnArgLeuPheTyrArgLysSerValTyrPheLeu 574
607 AGCACATTCAGAGACAGAGCTCTTCTTCTACCGTAAAGGTGTGTGGACAGCTG 663

SULT 4
972318
CUS 2M0246F07F Mouse 10kb plasmid UUGC2M library GSS 27-APR-2001
FINITION clone UUGC2M0246F07 F, genomic survey sequence.
SESSION AZ972318.1 GI:13843545
RSION AZ972318.1 GI:13843545
YWORDS Mus musculus (house mouse)
URCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 641)
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0246 row: F column: 07
Seq primer: CGTTGTAAACAGCCAGCCAGT
Class: plasmid ends
High quality sequence stop: 641.
Location/Qualifiers
1. .641
/organism="Mus musculus"

```

```

/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0246F07"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 155 a 195 c 157 g 134 t
ORIGIN
Alignment Scores: 1.09e-85 Length: 641
Pred. No.: 1082.00 Matches: 203
Score: 100.00% Conservative: 2
Percent Similarity: 99.02% Mismatches: 0
Best Local Similarity: 18.34% Indels: 0
Query Match: 28 Gaps: 0
DB:
US-09-042-460-2 (1-1122) x AZ972318 (1-641)

```

```

QY 244 AlAlYsLysAlaArgCysTyrProValProArgValGluGluGlyProHisArgGlnVal 263
Db 26 TCGCAGAAGGCCAGATCTATCTCTCCGAGAGTGTGAGGAGGAGCCACAGGCGAGTG 85
QY 264 LeuProThrProSerGlyLysSerTyrValProSerProAlaArgSerProGlnValPro 283
Db 86 CTACCAACCCCATCAGCAAAATCATGGTGTCCAAAGTCTCGTCCGTCGCCAGGTCCT 145
QY 284 ThrAlaGluLysAspLeuSerSerLysGlyLysValSerAspLeuSerLeuSerGlySer 303
Db 146 ACTGCAGNAAAGATTGTTCTTCTTAAGGAAGGTGTCTGACCTGAGTCTCTCTGGTGC 205
QY 304 ValCysCysLysHisLysProSerSerThrSerLeuLeuSerProProArgGlnAsnAla 323
Db 206 GTGTGCTGTAAACAACAAGCCAGCTCCACATCTCTGTGTCAACCCCGCCAAATGCC 265
QY 324 PheGlnLeuArgProPheIleGluThrArgHisPheLeuTyrSerArgGlyAspGlyGln 343
Db 266 TTTGAGTTCAGGCCATTTATTTAGACCAACGACATTTCTTTACTTCCAGGGAGATGGCCA 325
QY 344 GluArgLeuAsnProSerPheLeuLeuSerAsnLeuGlnProAsnLeuThrGlyAlaArg 363
Db 326 GAGCGTCTAAACCCCTCATTTCTTACTCTCAGCACTCAGCACTTGAATGGGCGCAG 385
QY 364 ArgLeuValGluIleIlePheLeuGlySerArgProArgThrSerGlyProLeuCysArg 383
Db 386 AGACTGGTGGAGATCATCTTTCTGGGCTCAAGGCCATGAGACATCAGGACCATCTCAGG 445
QY 384 ThrHisArgLeuSerArgArgTyrTyrGlnMetArgProLeuPheGlnGlnLeuLeuVal 403
Db 446 ACACACCGCTCATCGCTCGATCTGGCAGATCGGCCCTGTGTTCCACAGCTGCTGGTG 505
QY 404 AsnHisAlaGluCysGlnTyrValArgLeuLeuArgSerHisCysArgPheArgThrAla 423

```





Fax: 81-45-503-9216  
 Email: genome-res@gsic.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. *Genome Res.* 10 (10): 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11): 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. *Genome Res.* 11 (2): 281-289 (2001)  
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

## e mouse tissues.

## UTURES Location/Qualifiers

1. .599  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="5730412M20"  
 /sex="mixed"  
 /dev\_stages="8 days embryo"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 8 days embryo"  
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGGATCCAGAGGCTCTTTTCTTTTCTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of subtraction to  
 Ret = 100.0 Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGAGATTCGAGTTAATTAATTCGCCCCCCCC 3']. cDNA  
 was cloned into the XhoI and BamHI sites. Vector: a  
 modified pBluescript KS(+) after bulk excision from Lambda  
 FLIC i. Cloning sites, 5' end: Sali; 3' end: BamHI."  
 SE COUNT 91 a 185 c 161 g 142 t

## ALIGN

ignment Scores:  
 ed. No.: 2 61e-64 Length: 599  
 ore: 841.00 Matches: 161  
 cent Similarity: 100.00% Conservative: 0  
 st Local Similarity: 100.00% Mismatches: 0  
 ery Match: 14.25% Indels: 0  
 : 10 Gaps: 0

-09-042-460-2 (1-1122) x BB618671 (1-599)

1 MetThArgAlaProHArgCysProAlaValArgSerLeuLeuArgSerArgTyrArgGlu 20  
 115 ATGACCCGCGCTCTCTGTCGCCCGCGGTCCTCTGCTGCGAGCCGATACCGGGAG 174

Qy 21 ValTrpProLeuAlaThrPheValArgLeuGlyProGluGlyArgLeuValGln 40  
 Db 175 GTGTGGCCGCTGGCAACCTTTGTGGCGCCCTGGGGCCGAGGCGGCTGTGCA 234  
 Qy 41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMethHisTrp 60  
 Db 235 CCGGGGAGCCCGAAGATCTACCGCACTTTGGTTGGTGGCCCAATGCTAGTGTGCTGCTGG 294  
 Qy 61 GlySerClnProProAlaLeuSerPheHisGlnValSerSerLeuLysGluLeu 80  
 Db 295 GGCTCACAGCCTCCACCTGCCACCTTTCTTCCACAGGTGTCTATCCCTGAAGAGCTG 354  
 Qy 81 ValAlaArgValValGlnArgLeuCysGluArgAsnGluArgAsnValLeuAlaPheGly 100  
 Db 355 GTGGCCAGGGTTGTGCAGAGACTCTGCGAGCGCAACGAGAGAAAGCTGTGGCTTTTGGC 414  
 Qy 101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120  
 Db 415 TTTGAGCTGCTTAACGAGCGCAGAGCGGGCTCCCATGGCTTCCTAGTAGCGTGGT 474  
 Qy 121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140  
 Db 475 AGCTACTTGGCCCAACTCTTTATTGAGACCCCTGGCTGCTGAGTGGTGCATGGTACTG 534  
 Qy 141 LeuSerArgValGlyAspAspLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeu 160  
 Db 535 TTGAGCGAGTGGCGAGCAGCCTGCTGGTCTACCTGCTGGCACACTGTGCTTTATCTT 594  
 Qy 161 Leu 161  
 Db 595 CTG 597

## RESULT 7

BE371943

LOCUS

DEFINITION

601217728F1 NCI\_CGAP\_Lu29 Mus musculus cDNA clone IMAGE:3586614 5',

mRNA sequence.

ACCESSION

BE371943

VERSION

BE371943.1

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 (Bases 1 to 866)

AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM8748 row: h column: 07

High quality sequence stop: 639.

Location/Qualifiers

1. .866

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:3586614"

/tissue\_type="spontaneous tumor, metastatic to mammary."

Stem cell Origin."

/lab host="DH10B"

/clone lib="NCI\_CGAP Lu29"

/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: Sali;

Site 2: NotI; Cloned unidirectionally; Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

## FEATURES

source



```

41 ProGlyAspProLysIleTyrArgThrLeuValAlaGlnCysLeuValCysMetHisTrp 60
235 CCGGGGACCCGAGAGATCTACCGCATTTGGTTGCCCAATGCCCTAGTGTGCATGCACCTGG 294

61 GlySerGlnProProAlaAspLeuSerPheHisGlnValSerSerLeuLysGluLeu 80
295 GGTCTCAGAGCTTCCACCTGCGACCTTCTTCCACCGAGGTGTCTCCCTGGAAGAGCTG 354

81 ValAlaArgValValGlnArgLeuCysLysArgAsnGluArgAsnValLeuAlaPheGly 100
355 GTGGCCAGGGTGTGTGACAGACTCTGGAGCGCAACGAGAGAAACGCTGCTGCTTTGGC 414

101 PheGluLeuLeuAsnGluAlaArgGlyGlyProProMetAlaPheThrSerSerValArg 120
415 TTGAGCTGCTTAAAGAGCCAGAGCGGGGCTCCCATGGCTTCACTAGTACGGTGGCT 474

121 SerTyrLeuProAsnThrValIleGluThrLeuArgValSerGlyAlaTrpMetLeuLeu 140
475 AGCTACTTGGCCACACACTGTTATTGAGACCCCTGCTGTCTAGTGTGCATGGTACTTG 534

141 LeuSerArgValGlyAspAspLeuLeu---ValTyrLeuLeuAlaHisCysAlaLeuTy 159
535 TTGAGCCGAGTGGGCGGACGACTGCTGGGTCTAACTGCTGGGCACACTGGGGCTCCTT 594

159 rLeuLeuVal 162
595 TATCCTCTCTG 604

```

```

347 bp mRNA linear EST 25-JAN-2002
BR END06B09 Bain Rancourt retinoic acid induced ES cell neural
differentiation subtraction library Mus musculus cDNA clone 06B09
similar to gb|AF073311|AF073311 Mus musculus telomerase catalytic
subunit mRNA, complete cds, mRNA sequence.

```

```

AW244516
AW244516.1 GI:8051265
EST.
Mus musculus (house mouse)
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bain, G., Mansergh, F.C., Wride, M.A., Hance, J.E., Isogawa, A.,
Rancourt, S.L., Ray, W.J., Yoshimura, Y., Tsuzuki, T., Gottlieb, D.I.
and Rancourt, D.E.
ES cell neural differentiation reveals a substantial number of
novel ESTs
Funct. Integr. Genomics 1 (2), 127-139 (2000)
21652683
11793228
Contact: Rancourt DE
Department of Biochemistry and Molecular Biology
University of Calgary
3330 Hospital Drive N.W., Calgary, Alberta, T2N 4N1, Canada
Tel: 403 220 2888
Fax: 403 283 8727
Email: rancourt@calgary.ca; URL:http://www.acs.ucalgary.ca/
rancourt
DNA sequencing by: University Core DNA Services, University of
Calgary. Submitted sequence has been trimmed at both ends to remove
the adaptor oligos containing the EcoRI sites: i.e. GAATTCGACTA
(beginning) and TAGTCCGATTC (end) removed. Therefore, reported
insert length is longer than actual EST sequence length.
Insert Length: 359 Std Error: 10.00
Seq primer: T3 Or T7.
Location/Qualifiers
1. 347
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129"
/db_xref="taxon:10090"

```

FEATURES  
source

```

/clone="06B09"
/cell_type="embryonic stem (ES) cell"
/cell_line="D3"
/lab_host="DH5 alpha"
/clone_lib="Bain Rancourt retinoic acid induced ES cell
neural differentiation subtraction library"
/notes="Vector: pBluescript II SK+ (Stratagene); Site 1:
EcoRI; Site 2: EcoRI; Library constructed by Dr. Gerard
Bain (present address: Hoechst-ARIAD Genomic Center,
ARIAD Pharmaceuticals Inc., 26 Landsdowne Street,
Cambridge, Massachusetts, 02139-4234, U.S.A.). To isolate
cDNAs corresponding to mRNAs which are upregulated during
the neural differentiation of ES cells in vitro, the
subtractive hybridization technique of Wang and Brown [1]
was employed. Poly(A)+ RNA was prepared from both
undifferentiated ES cells and from embryoid bodies which
had been cultured for 4 days in the absence of RA followed
by an additional 3 days in the presence of 0.5 (M RA
(4-/3+ cells). These poly(A)+ RNAs were converted to
double-stranded cDNA using the Superscript Choice System
(Gibco). Aliquots of both cDNAs were digested with the
restriction enzymes AluI and AluI plus Real. An adaptor
oligo [1] containing an EcoRI site was ligated to the ends
of the restricted cDNAs to provide primer binding sites
and large amounts of each cDNA population were then
produced by the polymerase chain reaction (PCR) as
described [1]. Amplified cDNA from undifferentiated ES
cells was biotinylated using Photoprobe biotin (Vector
Laboratories) according to the manufacturer's protocol.
2.5 ug of amplified cDNA from 4-/3+ cells was mixed with
50 ug of biotinylated ES cell cDNA, denatured by boiling,
and hybridized for 20 h. Double stranded cDNAs containing
biotin were removed by streptavidin/phenol treatment as
described [1]. The remaining subtracted cDNA was mixed
with an additional 25 mg of biotinylated ES cell cDNA,
denatured by boiling, and hybridized for 2 h. The
streptavidin/phenol treatment was repeated and the
remaining cDNA was amplified by PCR [Wang and Brown, 1991
]. Two additional rounds of subtraction were repeated
exactly as described above. The cDNA obtained from this
subtraction procedure was digested with EcoRI and ligated
to pBS II SK+ (Stratagene) followed by transformation
into E. coli DH5 cells. Individual colonies were picked
and the corresponding plasmids were isolated either by an
alkaline lysis miniprep procedure [2], or using the
Qiaprep spin miniprep kit (Qiagen). Sequence analysis was
performed using the Big Dye Cycle Sequencing kit and an
ABI373 sequencer (University Core DNA Services,
University of Calgary). 1. Wang, Z; Brown, D.D (1991) A gene
expression screen. Proc. Natl. Acad. Sc i. USA 88,
11505-11509. 2. Sambrook, J; Fritsch, E.F; Maniatis, T. (1989)
Molecular Cloning: A Laboratory Manual. Cold Spring Harbor
Laboratory Press, Cold Spring Harbor, New York."

```

BASE COUNT 78 a 88 c 100 g 81 t  
ORIGIN

Alignment Scores:  
Pred. No.: Length: 347  
Score: 595.00 Matches: 115  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.08% Indels: 0  
DB: 9 Gaps: 0

US-09-042-460-2 (1-1122) x AW244516 (1-347)

QY 731 CysIleArgGlnTyrAlaValValArgAspSerGlnGlyGlnValHisLysSerPhe 750  
Db 345 TGTATCCGCCAGTATGCACTGGTCCGAGAGATAGCCAGGCCCAAGTCCACAGTCTCTTT 286  
QY 751 ArgArgGlnValThrThrLeuSerAspLeuGlnProTyrMetGlyGlnPheLeuTyHis 770  
Db 285 AGGAGACAGGTACACACCTCTCTGACCTCCAGCCATACATGGGCCAGTTCCTTAGCAT 226

771 LeuGlnAspSerAspAlaSerAlaLeuArgAsnSerValValIleGlnSerIleSer 790  
 225 CTGAGGATTTCAGATGCGAGTGCAGTCCGTTGTTCATGAGCAGAGCATCTCT 166  
 791 MetAsnGluSerSerSerSerLeuPheAspPhePheLeuHisPheLeuArgHisSerVal 810  
 165 ATGAATGAGAGCAGCAGCAGCTGTTTTCCTTCTTCACCTCTCTGCGTCCAGATGTC 106  
 811 ValIleGlnGlyAspArgCysThrGlnCysGlnGlyIleProGlnGlySerSerLeu 830  
 105 GTAAAGATTGCTGACAGTGTATACGACATGCCAGGCGATCCCCAGGCGCTCCAGCCTA 46  
 831 SerThrLeuLeuCysSerLeuCysPheGlyAspMetGluAsnLys 845  
 45 TCCACCTGCTCTGCACTCTGTGTTTCGGAGACATGGAGACAAAG 1

UT 10  
 24748  
 US  
 INITIATION K-EST0096335 S22SNU16n1 Homo sapiens cDNA clone S22SNU16n1-99-E07  
 5', mRNA sequence.  
 ESSION BM824748  
 STION BM824748.1 GI:19181161  
 WORDS EST.  
 RGANISM Homo sapiens (human)  
 RCE Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 EREFERENCE  
 UTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 UNPUBLISHED  
 CONTACT: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eosun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 99 row: E column: 07  
 High quality sequence stop: 492.  
 Location/Qualifiers  
 1. 492  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S22SNU16n1-99-E07"  
 /sex="F"  
 /tissue\_type="Ascites"  
 /cell\_type="lymphoblast-like"  
 /cell\_line="SNU-16"  
 /lab\_host="DH10B"  
 /clone\_lib="S22SNU16n1"  
 /note="Organ: Stomach; Vector: pT7T3-Pac; Site: 1. EcoRI;  
 Site 2: NotI; The S22SNU16 library was contributed by the  
 Soares laboratory and it was constructed as described by  
 Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome  
 Research 6(9): 791-806. RNA was prepared from harvested  
 cells of SNU-16 culture. SNU-16 cell was obtained from  
 Korean Cell Line Bank (KCLB). SNU-16 was established from  
 ascitic fluids of Korean patients by Park J.G. et al.  
 (1990), Cancer Res 50: 2773-2780."

IE COUNT 96 a 152 c 131 g 113 t

GIN

gment Scores:  
 id. No.: 2.77e-39 Length: 492  
 re: 558.00 Matches: 109  
 cent Similarity: 80.27% Conservative: 9  
 it Local Similarity: 74.15% Mismatches: 29

Query Match: 9.46% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-042-460-2 (1-1122) x BM824748 (1-492)  
 QY 898 ValAsnPheProValGluProGlyThrLeuGlyGlyAlaAlaProTyGlnLeuProAla 917  
 DB 1 GTGAATCTCCCTGTAGAAGACGAGGCCCTGGGTGGCGCTTTTGTTCAGATGCCGCC 60  
 QY 918 HisCysLeuPheProTyrCysGlyLeuLeuLeuAspThrGlnThrLeuGluValPheCys 937  
 DB 61 CACGGCTATTCCCTGGTGGCGCTGTCTGTGGATACCCGACCCCTGGAGGTGCAGAGC 120  
 QY 938 AspTyrSerGlyTyrAlaGlnThrSerIleTyrThrSerLeuThrPheGlnSerValPhe 957  
 DB 121 GACTACTCCAGTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCACCGGGCTTC 180  
 QY 958 LysAlaGlyIleThrMetArgAsnLysLeuLeuSerValLeuArgLeuLysCysHisGly 977  
 DB 181 AAGGCTGGGAGGAACATGCGTCGCAACTCTTTGGGTCTTGGCGCTGAAGTGTCCAGC 240  
 QY 978 LeuPheLeuAspLeuGlnValAsnSerLeuGlnThrValCysIleAsnIleTyrLysIle 997  
 DB 241 CTGTTTCTGGATTTCAGGTGAACAGCTCCAGAGCGGTGCACCAACATCTACAGATC 300  
 QY 998 PheLeuLeuGlnAlaTyrArgPheHisAlaCysValIleGlnLeuProPheAspGlnArg 1017  
 DB 301 CTCCTGCTGCAGCGGTACAGGTTTCAGCATGTGTGTGCAGCTCCATTTTCATCAGCAA 360  
 QY 1018 ValArgLysAsnLeuThrPhePheLeuGlyIleIleSerSerGlnAlaSerCysCysTyr 1037  
 DB 361 GTTTGAAAGAACCCCACTTTCTTCGCGGTCACTCTGACACGGGCTTCCTCTGTCTAC 420  
 QY 1038 AlaIleLeuLysValLysAsn 1044  
 DB 421 TCCATCTCGAAGCCCAAGAAC 441  
 RESULT 11  
 BU452535  
 LOCUS 603767927F1 CSQRN14 Gallus gallus cDNA clone CHEST696f7 5', mRNA  
 DEFINITION  
 sequence.  
 ACCSSION BU452535  
 VERSION BU452535.1 GI:25941846  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Gallus gallus (chicken)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 753)  
 Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,  
 Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.U.  
 A Comprehensive Collection of Chicken cDNAs  
 Curr. Biol. 12 (22), 1965-1969 (2002)  
 22335534  
 MEDLINE  
 PUBMED  
 12445392  
 COMMENT  
 Contact: Simon Hubbard  
 Department of Biomolecular Sciences  
 University of Manchester Institute of Science and Technology (UMIST  
 )  
 PO Box 88, Manchester, M60 1QD, UK  
 Tel: 01612008930  
 Fax: 01612360409  
 Email: Simon.Hubbard@umist.ac.uk.  
 Location/Qualifiers  
 1. 753  
 /organism="Gallus gallus"  
 /mol\_type="mRNA"  
 /strain="Layer"  
 /db\_xref="taxon:9031"  
 /clone="CHEST696f7"  
 /sex="Female"  
 FEATURES  
 source

```
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRBN14"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following ligation in the first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
E COUNT      230 a 155 c 132 g 236 t
GIN

gment Scores:
d. No.:      1.98e-37      Length:      753
re:          540.50      Matches:      106
cent Similarity: 76.36%      Conservative: 20
t Local Similarity: 64.24%      Mismatches: 36
ry Match:     9.16%      Indels:      3
              13      Gaps:      1

09-042-460-2 (1-1122) x BU452535 (1-753)
961 LysThrMetArgAnSerLeuLeuSerValLeuArgLeuLysCysHisGlyLeuPheLeu 980
5 AAAAAACATGAATGCAATGACGTGACGCTCCTCAACTGAATGCCATCCTTTACTTCT 64
981 AspLeuGlnValAnSerLeuGlnThrValCysLeuAsnLeuLysLeuPheLeuLeu 1000
65 GACITTAAGATCAACAGCCTTCACAGAGTCTCTTAATTAACATCTACAGATATTTTACTT 124
1001 GlnAlaTyrArgPheHisAlaCysValLeuGlnLeuProPheAspGlnArgValArgLys 1020
125 CAGGCTTACAGTTCCATGCTGCTGTGTTCTTTCAGCTTCCATTCACCAAGAGTTAGGAAT 184
1021 AnLeuThrPheLeuGlyLeuLeuSerSerGlnAlaSerCysCysTyrAlaLeuLeu 1040
185 AATCCTGATTTCTTCTAAGATCATCTCTGATACATCTGCTTCATGCTGCTATTTTATCTG 244
1041 LysValLysAnProGlyMetThrLeu-----LysAlaSerGlySerPheProPro 1057
245 AAAGCTAAATCCAGAGTTCTTCTAGTAGCAAGATCCTCTGATGTTCCCTTTT 304
1058 GluAlaAlaHisTyrLeuCysTyrGlnAlaPheLeuLysLeuAlaHisSerVal 1077
305 GAGGCAGCAGAAATGGCTGTGTACCATGCTCCCTTCATTCACAACTGTCACCAACCAAGTT 364
1078 IleTyrLysCysLeuLeuGlyProLeuArgThrAlaGlnLysLeuLeuCysArgLysLeu 1097
365 ATTACAAATGCTTACTTAAGCCCTTAAGTCTAAGATGATGCTGTTTGGAGATC 424
1098 ProGluAlaThrMetThrIleLeuLeuAlaAlaAspProAlaLeuSerThrAspPhe 1117
425 CCAGGGATACATATGAACCTGCTGAAGCGTGAAGCAACCATCGCTTGTCTCAAGATTTC 484
1118 GlnThrIleLeuAsp 1122
485 AAAACTACTATGGAC 499

FEATURES
source
location/Qualifiers
1..326
/organism="Mus musculus"

KEYWORDS
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 326)
OKazaki.Y., Furuno.M., Kasukawa.T., Adachi.J., Bono.H., Kondo.S.,
Nikado.I., Otsu.N., Saito.R., Suzuki.H., Yamanaka.I., Kiyosawa.H.,
Yagi.K., Tomaru.Y., Hasegawa.Y., Nogami.A., Schonbach.C.,
Gojobori.T., Baldarelli.R., Hill.D.P., Bult.C., Hume.D.A.,
Quackenbush.J., Schriml.L.M., Kanapin.A., Matsuda.H., Batalov.S.,
Beisel.K.M., Blake.J.A., Bradt.D., Brusic.V., Chothia.C., Corbani
L.E., Cousins.S., Dalla.E., Dragani.T.A., Fletcher.C.F., Forrest
.A., Frazer.K.S., Gaasterland.T., Gariboldi.M., Gissi.C., Godzik.A.,
Gough.J., Grimmond.S., Gustincich.S., Hirokawa.N., Jackson.I.J.,
Jarvis.E., Kanai.A., Kawaji.H., Kawasawa.Y., Kedzierski.R.M.,
King.B.L., Konagaya.A., Kurochkin.I.V., Lee.Y., Lenhard.B., Lyons
.P.A., Maglott.D.R., Maltais.L., Marchionni.L., McKenzie.L., Miki
.H., Nagashima.T., Numata.K., Okido.T., Pavan.W.J., Pertea.G.,
Pesole.G., Petrovsky.N., Pillai.R., Pontius.J.U., Qi.D.,
Ramachandran.S., Ravasi.T., Reed.J.C., Reed.D.J., Reid.J., Ring
.B.Z., Ringwald.M., Sandelin.A., Schneider.C., Sempile.C.A., Setou
.M., Shimada.K., Sultana.R., Takenaka.Y., Taylor.W.S., Teasdale
.R.D., Tomita.M., Verardo.R., Wagner.L., Wahlstedt.C., Wang.Y.,
Watanabe.Y., Wells.C., Wilming.L.G., Wynshaw-Boris.A., Yaginuma
.M., Yang.I., Yang.L., Yuan.Z., Zavolan.M., Zhu.Y., Zimmer.A.,
Carninci.P., Hayatsu.N., Hirozane-Kishikawa.T., Konno.H., Nakamura
.M., Sakazume.N., Sato.K., Shiraki.T., Waki.K., Kawai.J., Aizawa.K.,
Arakawa.T., Fukuda.S., Hara.A., Hashizume.W., Imotani.K., Ishii
.Y., Itoh.M., Kagawa.I., Miyazaki.A., Sakai.K., Sasaki.D., Shibata
.K., Shinagawa.A., Yasunishi.A., Yoshino.M., Waterston.R., Lander
.E.S., Rogers.J., Birney.E. and Hayashizaki.Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
22354683
12466851
Contact: Yoshhide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL: http://genome-res@gs.riken.go.jp/
Aizawa.K., Akimura.T., Arakawa.T., Carninci.P., Fukuda.S., Hirozane
.T., Imotani.K., Ishii.Y., Itoh.M., Kawai.J., Konno.H., Miyazaki.A.,
Murata.M., Nakamura.M., Nomura.K., Numazaki.R., Ohno.M., Sakai.K.,
Sakazume.N., Sasaki.D., Sato.K., Shibata.K., Shiraki.T., Tagami
.M., Waki.K., Watanaka.A., Muramatsu.M. and Hayashizaki.Y. Direct
Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
location/Qualifiers
1..326
/organism="Mus musculus"
```

```

/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="1930261124"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"
E COUNT      58 a  92 c  98 g  78 t
GIN

gment Scores:
d. No.:      2.15e-35      Length:      326
re:          511.00      Matches:      103
cent Similarity: 99.04%      Conservative: 0
t Local Similarity: 99.04%      Mismatches: 0
ry Match:      8.66%      Indels:      1
          13      Gaps:      0

```

09-042-460-2 (1-1122) x BY149368 (1-326)

```

67 AlaAspLeuSerPheHisGlnValSerSerLeuLeuValAlaArgValValGln 86
2 GCCACCTTTCTTCCACAGGTGTCATCCCTGAAAGAGCTGCTGCCAGGTTTGTGCAG 61
87 ArgLeuCysGluArgAenGluAArgAenValLeuAlaPheGlyPheGluLeuLeuAenGlu 106
62 AGACTCTGCCAGCGCACGAGAGAAAGCTGCTGGCTTTTGGCTTTGAGCTGCTTAACGAG 121
107 AlaArgGlyGlyProMetAlaPheThrSerSerValArgSerTyrLeu-ProAsnTh 126
122 GCCAGAGCGCGGCTCCCACTAGGCGCTTCACTAGTAGGCTGCTAGTACTTGCCTCCCAACAC 181
126 rValleuLeuThrLeuArgValSerGlyAlaTrpMetLeuLeuLeuSerArgValGlyAs 146
182 TGTTATTGAGACCTTGGCTGCTAGTGTGATGATGCTACTGTTGAGCGGAGTGGCGA 241
146 pAspLeuLeuValTyrLeuLeuAlaHisCysAlaLeuTyrLeuLeuValProSerCy 166
242 CGACTCTGCTGCTACCTGCTGCGACACTGCTCTTTATCTTCGTGTCGCCCCAGCTG 301
166 SaluTyrGln 169
302 TGCCTACCAG 311

```

ULT 13  
88013

US BI388013 775 bp mRNA linear EST 20-MAY-2003  
INITIATION BFL26\_002716 Amphioxus 26hr cdna library (Name convention: BFL26 or  
MFMGP531) Branchiostoma floridae cdna clone MFMGP531A0937 5', mRNA  
sequence.

ESSION BI388013  
STON BI388013.1 GI:30922834  
WORDS EST.

RGANISM Branchiostoma floridae (Florida lancelet)  
Branchiostoma floridae

Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.

ERENCE 1 (bases 1 to 775)  
UTHORS Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A., Herwig  
, R., Vingron, M. and Lehrach, H.

TITLE New evidence for genome wide duplications at the origin of  
vertebrates using an amphioxus gene set and completed animal

Genomes Res. 13 (6), (2003) In press  
JOURNAL Genome Res.

MENT Laboratory 145, dept. Lehrach  
Contact: Panopoulou G

Max-Planck-Institut fuer Molekulare Genetik  
Inherst-63-73, D-14195 Berlin, Germany

Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting

(ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/tfide> )amphioxus/main.html Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTACACTTTATGTTCCGGCTCG 3' (M13BSP)

BACKWARD: 5' GCTATTAGCAGCTGCGAAAGGGGATGTG 3' (M13PSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGCTCCGGAATTCCTCCGGT-3' pSPORT3/86

High quality sequence stop: 775.

#### FEATURES

Location/Qualifiers

1..775

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MFMGP531A0937"

/tissue\_type="whole embryo"

/dev\_stage="26 hrs (neurula stage)"

/lab\_host="E. coli, XL1 blue"

/clone\_lib="Amphioxus 26hr cdna library (Name convention: BFL26 or MFMGP531)"

/note="Vector: pSPORT1 (Gibco BRL); Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSPORT1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGGCGGCC (T)15-3' and a SalI 5'-TCGACCACTGCTCCG-3' adapters (Gibco BRL)."

BASE COUNT 176 a 217 c 194 g 186 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 3,14e-33 Length: 775

Score: 493.50 Matches: 110

Percent Similarity: 56.47% Conservative: 34

Best Local Similarity: 43.14% Mismatches: 104

Query Match: 8.36% Indels: 7

DB: 12 Gaps: 3

US-09-042-460-2 (1-1122) x BI388013 (1-775)

Qy 810 valVallylleGlyAspArgCysTyrThrGlnCysGlnGlylleProGlnGlySerSer 829  
Db 24 GTGATCCAGATCGGGGGAGCTACTACCGGCAGATCAGCGGGATCAGTCAGGGATCTCTGC 83  
Qy 830 LeuSerThrLeuLeuCysSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 849  
Db 84 CTGTCCACCTCTCTGTAGTCTTCTACGCTAACATGAGAGAGAGACATCTCTGTGGC 143  
Qy 850 valGlnArgAspGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 869  
Db 144 ATGACATCATGATGGGTGCTGCTGCTGTGGTGGATGCTTCTCTGTGGTGGATGCTTCT 203  
Qy 870 LeuAspGlnAlaLysThrPheLeuSerThrLeuValHisGlyValProGluTyrGlyCys 889  
Db 204 CTGGACACGCTACAGCTTCTCTGTACCATGTTGGATGTTGATCCCTGATATGGGTGC 263  
Qy 890 MetIleAsnLeuGlnLysThrValValAsnPheProValGluProGlyThrLeuGlyGly 909  
Db 264 AGCGTTTCATCCGACAAAGTCTGACCAACTTCTCCCTGCCAGCAC-----AGGAAT 314  
Qy 910 AlaAlaProTyrGlnLeuProAlaHisCysLeuPheProTyrCysGlyLeuLeuLeuLeu 929  
Db 315 GCTGTGGTGAATGTCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 374  
Qy 930 ThrGlnThrLeuGluValPheCysAspTyrSerGlyTyrAlaGlnThrSerIleLeuThr 949





TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgapbs-re@mail.nih.gov  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 2187 Std Error: 0.00  
 Seq primer: -28m13 rev2 ET from Amersham  
 High quality sequence stop: 385.  
 Location/Qualifiers  
 1. 389

Search completed: February 1, 2004, 19:29:33  
 Job time : 5272 secs

# FEATURES

source  
 1. 389  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:712562"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP GCB1"  
 /note="Vector: p773D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD20+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 [5'-TGTTACATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'  
 ]. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified p773 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 E COUNT 87 a 102 C 123 g 77 t  
 GIN

## Alignment Scores:

d. NO.: 5.3e-31 Length: 389  
 re: 463.00 Matches: 87  
 cent Similarity: 82.03% Conservative: 18  
 t Local Similarity: 67.97% Mismatches: 23  
 ry Match: 7.85% Indels: 0  
 Gaps: 0

09-042-460-2 (1-1122) x AA281296 (1-389)

532 AlaThrPheLeuPheTrpLeuMetAspThrTyrValValGlnLeuLeuArgSerPhePhe 551  
 1 GCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCTGTCGAGCTGCTCAGGCTTCTTT 60  
 552 TyrIleThrGluSerThrPheGlnLysAsnArgLeuPhePheTyrArgLysSerValTrp 571  
 61 TATGTTCAGGAGACCAAGTTTCAAAAGAACAGGCTCTTTTCTACCGAAGAGTGTCTGG 120  
 572 SerLysLeuGlnSerIleGlyValArgGlnHisLeuGluArgValArgLeuArgGlnLeu 591  
 121 AGCAAGTTTCAAGCATTTGGAATCAGACAGACACTTGAAGAGGGTGCAGCTGCGGACGTG 180  
 592 SerGlnGluGluValArgHisHisGlnAspThrTrpLeuAlaMetProIleCysArgLeu 611  
 181 TCGGAAGCAGAGGTTCAGGCAGCATTCGGGAAGCCAGCGCCCTGCTGACGTCCAGACTC 240  
 612 ArgPheIleProLysProAsnGlyLeuArgProIleValAsnMetSerTyrSerMetGly 631  
 241 CGCTTCATCCCAAGCTTGACGGGCTCGGCCGCAATTGTGACATGACTACGTCTGTTGGA 300  
 632 ThrArgAlaLeuGlyArgArgLysGlnAlaGlnHisPheThrGlnArgLeuLysThrLeu 651  
 301 GCCAGACGTTCCGCAGAGAAAGAGGCGCGAGCGTCTCACCTCGAGGGTGAAGGCACTG 360  
 652 PheSerMetLeuAsnTyrGluArg 659  
 361 TTCAGCGTGTCAACTACGAGCG 384

Search completed: February 1, 2004, 19:29:33  
 Job time : 5272 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: February 2, 2004, 09:12:03 ; Search time 41 Seconds

(without alignments)  
933.003 Million cell updates/sec

le: US-09-780-669-114

ue: 1 MQCFSPFKTWMLFNLLIFL.....IGGLELAAMVSMYLYCNLQ 241

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : A\_Geneseq\_19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	1258	100.0	241	19 AAW71870	Amino acid encoded
2	1258	100.0	241	19 AAW59954	Amino acid sequenc
3	1258	100.0	241	19 AAW59386	Prostate tumour sp
4	1258	100.0	241	19 AAW58380	Human secreted pro
5	1258	100.0	241	21 AAY82003	Human immunogenic
6	1258	100.0	241	21 AAU59764	Human prostate CDN
7	1258	100.0	241	22 AAU04962	Human prostate tum
8	1258	100.0	241	22 AAM01118	Human prostate-spe
9	1258	100.0	241	22 AAG99003	Human prostate-spe

10	1258	100.0	241	22 AAB74801	Prostate tumour an
11	1258	100.0	241	22 AAB90679	Human AK415_4 prot
12	1258	100.0	241	23 ABG94412	Human prostate tum
13	1258	100.0	241	23 ABG76666	Prostate tumour pr
14	1258	100.0	241	23 ABG95223	Human N1-1862 prot
15	1258	100.0	241	24 ABU71654	Prostate cancer as
16	1254	99.7	241	23 ABG96434	Human ovarian canc
17	1253	99.6	240	22 AAG62153	Human P503S invent
18	1241.5	98.7	240	23 ABB89113	Human polypeptide
19	1209	96.1	233	19 AAW75060	Human secreted pro
20	1199	95.3	273	19 AAW61618	Clone HPWAB25 of T
21	1182	94.0	258	23 ABG96433	Human ovarian canc
22	911	72.4	241	23 ABB09512	Human TSPAN-1-like
23	575	45.7	126	20 AAY35991	Extended human sec
24	526	41.8	121	23 ABP43214	Human ovarian anti
25	480	38.2	101	20 AAY13159	Human secreted pro
26	473	37.6	97	20 AAY12381	Human 5' EST secre
27	453	36.0	99	19 AAW47275	Human HPK-1A C4.8
28	452	35.9	89	20 AAY11862	Human 5' EST secre
29	407	32.4	103	23 ABP32558	Human glycoprotein
30	382	30.4	140	20 AAW86331	Kidney injury asso
31	373	29.7	108	20 AAY11883	Human 5' EST secre
32	368.5	29.3	248	22 AAB93282	Human protein sequ
33	368.5	29.3	248	22 AAB49503	Clone HCEIK90 #1.
34	368.5	29.3	248	23 AAB64749	Human protein SEQ
35	368.5	29.3	248	23 AAW79168	Human CD-81-like p
36	368.5	29.3	248	23 AAW48320	Human tetraspan.
37	361	28.7	249	23 AAW77412	Human NOV6 protein
38	351.5	27.9	236	22 AAB88457	Human membrane or
39	330.5	26.3	247	22 AAW47215	Human NOV7 protein
40	310	24.6	239	22 AAW93980	Human stomach canc
41	310	24.6	239	22 AAB93885	Human protein sequ
42	297	23.6	238	19 AAW61622	Clone HTPBA27 of T
43	296	23.5	164	22 AAB49509	Clone HCEIK90 #2.
44	288	22.9	96	22 AAG75719	Human colon cancer
45	288	22.9	252	19 AAW61624	Clone HPFEK40 of T

#### ALIGNMENTS

#### RESULT 1

AAW71870  
ID AAW71870 standard; Protein; 241 AA.

XX AAW71870;

XX 06-JAN-1999 (first entry)

XX Amino acid encoded by prostate tumour clone N1-1862.

XX Prostate; cancer; tumour; vaccine; immunogen; clone.

XX Homo sapiens.

XX WO9837093-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US03492.

XX 09-FEB-1998; 98US-0020956.

XX 25-FEB-1997; 97US-0806039.

XX 01-AUG-1997; 97US-0904804.

XX (CORI-) CORIXA CORP.

XX Dillon DC, Xu J;

XX WPI; 1998-609886/51.

XX N-PSDB; AAW61202.

XX Polypeptides comprising immunogenic portions of prostate proteins -

used in a vaccine for the treatment of prostate cancer

Example 1; Page 84-85; 130pp; English.

The present sequence is an immunogenic portion of a prostate tumour protein. The immunogen, or the DNA encoding it, can be used as a vaccine for the treatment of prostate cancer. The immunogen was isolated from a prostate tumour cDNA library obtained by subtracting a prostate tumour cDNA expression library with a normal tissue cDNA library.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 61 IAAGVVVFALGFLGCGYAKTESKCALVTFFILLIFIAEVAALVAVVVTMAEHFLTL 120  
 61 IAAGVVVFALGFLGCGYAKTESKCALVTFFILLIFIAEVAALVAVVVTMAEHFLTL 120  
 121 LVVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 121 LVVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

SULT 2  
 W59954

AAW59954 standard; Protein; 241 AA.

AAW59954;

02-DEC-1998 (first entry)

Amino acid sequence of the human tumour-associated antigen.

Human; tumour-associated antigen; PRAT; stimulation; cell proliferation; antagonist; cancer; genetic defect; sickle cell anaemia; agonist; antibody; hybridisation; probe.

Homo sapiens.

WO9838310-A1.

03-SEP-1998.

27-FEB-1998; 98WO-US03953.

28-FEB-1997; 97US-0808148.

(INCY-) INCYTE PHARM INC.

Goli SK, Hillman JL;

WPI; 1998-481208/41.

N-PSDB; AAV54014.

Human tumour-associated antigen PRAT - useful for stimulating cell proliferation and screening for antagonists useful to treat or prevent cell proliferation disorders e.g. cancers

PS Disclosure; Fig 1A-1C; 54pp; English.

XX This is the amino acid sequence of the human tumour-associated antigen (PRAT) used in the method of the invention for stimulating cell proliferation and screening for antagonists useful to treat or prevent cell proliferation disorders such as cancer, and genetic defect e.g. sickle cell anaemia. The polypeptides can be combined with a suitable carrier in pharmaceutical compositions, and also used to screen for antagonists, agonists, and to generate antibodies. PRAT agonists can be added to a cell to stimulate cell proliferation by increasing or prolonging the activity of PRAT as above. The antagonists can be combined with a suitable carrier in pharmaceutical compositions, which can be administered to subjects to treat or prevent disorders associated with cell proliferation, especially cancers. Antibodies specific for PRAT may be used directly as antagonists, or indirectly as a targeting or delivery mechanism to bring pharmaceutical agents to PRAT-expressing cells. They are also useful to diagnose conditions or diseases characterised by PRAT expression and to monitor therapeutic interventions. The polynucleotide encoding PRAT, or complementary sequences, can be used to produce hybridisation probes, useful to detect polynucleotides or fragments encoding PRAT, e.g. to diagnose diseases relating to polypeptide expression or monitor PRAT regulation during therapeutic intervention.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 Db 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60  
 QY 61 IAAGVVVFALGFLGCGYAKTESKCALVTFFILLIFIAEVAALVAVVVTMAEHFLTL 120  
 Db 61 IAAGVVVFALGFLGCGYAKTESKCALVTFFILLIFIAEVAALVAVVVTMAEHFLTL 120  
 QY 121 LVVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 Db 121 LVVPAIKKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPPPFCNDNVT 180  
 QY 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 Db 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240  
 QY 241 Q 241  
 Db 241 Q 241

RESULT 3

AAW69386  
 ID AAW69386 standard; Protein; 241 AA.

XX AAW69386;

XX 25-MAR-2003 (updated)

DT 08-DEC-1998 (first entry)

XX Prostate tumour specific gene clone N1-1862 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection; therapy.

XX Homo sapiens.

XX WO9837418-A2.

XX 27-AUG-1998.

XX 25-FEB-1998; 98WO-US03690.

XX

25-FEB-1997; 97US-0806596.  
01-AUG-1997; 97US-0904809.  
09-FEB-1998; 98US-0020747.

(CORI-) CORIXA CORP.

Dillon DC, Xu J;

WPI; 1998-480805/41.  
N-PSDB; AAV58587.

Novel human prostate specific tumour protein and fragments - useful  
for detecting and treating prostate cancers

Example 1; Page 89-90; 141pp; English.

This sequence is encoded by a human prostate tumour specific gene, and  
can be used in the method of the invention. The method is for detecting  
prostate cancer comprises contacting a biological sample with an agent  
able to bind an immunogenic portion of a prostate protein (such as  
this protein sequence). An antibody which binds to an immunogenic  
portion of the prostate protein, and the method can be used to detect,  
monitor progression of, or treat prostate cancers. The antibody may  
also be conjugated to a therapeutic agent for use in therapy of prostate  
cancers.

(Updated on 25-MAR-2003 to correct PR field.)

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSTIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPLSSAMQFVNVGYFL 60

1 MQCFSTIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPLSSAMQFVNVGYFL 60

61 IAGVWVVFALGFLGCGYKATESKCALVTPFFILLIFIAEVAALVAVVYTTMAEHFTL 120

61 IAGVWVVFALGFLGCGYKATESKCALVTPFFILLIFIAEVAALVAVVYTTMAEHFTL 120

121 LVVPAIKDYGSQEDFTQVWNTTMMKLGKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

121 LVVPAIKDYGSQEDFTQVWNTTMMKLGKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

241 Q 241

241 Q 241

ULT 4

58380

AAW58380 standard; Protein; 241 AA.

AAW58380;

14-SEP-1998 (first entry)

Human secreted protein AR415\_4.

AR415\_4; secreted protein; protein factor; human.

Homo sapiens.

Key Location/Qualifiers

Peptide 14..26

/note= "putative leader/signal peptide or  
transmembrane domain"

Protein 27..241

/label= Mat\_protein

WO9817687-A2.

30-APR-1998.

24-OCT-1997; 97WO-US19590.

24-OCT-1997; 97US-0740274.

25-OCT-1996; 96US-0740274.

(GEM ) GENETICS INST INC.

Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;

Racie LA, Spaulding V, Treacy M;

WPI; 1998-261426/23.

N-PSDB; AAV30916.

Nucleic acid encoding secreted protein from human cells - useful,  
e.g. as immunomodulator, antitumour agent, promoters of tissue  
growth, haemostatic and thrombolytic agents etc.

Claim 8; Page 67-68; 114pp; English.

This polypeptide, designated AT415\_4, is a novel human secreted  
protein. Its amino acid sequence was deduced from a full-length  
AT415\_4 cDNA clone (see AAV30916) isolated from a human adult retina  
cDNA library. The predicted amino acid sequence shows homology to  
human AAM35252 and CO-029 tumour associated antigens, and computer  
predictions suggest a potential transmembrane domain centered  
around amino acid 100 of the protein. 11 Novel human secreted  
proteins (see AAW58580-90) are claimed. These can be expressed in  
recombinant host cells for analysis, characterisation, diagnostic  
or therapeutic use. They can also be used as tissue or mol.wt.  
markers, to generate antibodies, and in interaction trap assays.  
They may have biological activities, e.g. cytokine, immunomodulator,  
haematopoiesis regulating activity, tissue growth activity, activin  
or inhibin activity, chemotactic or chemokinetic activity,  
haemostatic and thrombolytic activity, receptor/ligand activity,  
antiinflammatory, cadherin and tumour invasion suppressor activity,  
and tumour inhibition activity. The proteins can be expressed in  
vivo from DNA, introduced in gene therapy vectors.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 19; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQCFSTIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPLSSAMQFVNVGYFL 60

Db 1 MQCFSTIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPLSSAMQFVNVGYFL 60

Qy 61 IAGVWVVFALGFLGCGYKATESKCALVTPFFILLIFIAEVAALVAVVYTTMAEHFTL 120

Db 61 IAGVWVVFALGFLGCGYKATESKCALVTPFFILLIFIAEVAALVAVVYTTMAEHFTL 120

Qy 121 LVVPAIKDYGSQEDFTQVWNTTMMKLGKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

Db 121 LVVPAIKDYGSQEDFTQVWNTTMMKLGKCCGFTNTDFEDSPYFKENSAPFPCCNDVNT 180

Qy 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

Db 181 NTANETCTKQKADQKVEGCFNQLLYDIRTNVTVGGVAAAGIGGLELAAMIVSMYLYCNL 240

Qy 241 Q 241

Db 241 Q 241

RESULT 5  
AAW582003

AAV82003 standard; Protein; 241 AA.

AAV82003;

13-JUN-2000 (first entry)

Human immunogenic prostate tumour protein sequence SEQ ID NO:114.

Human; prostate cancer; diagnosis; tumour; gene therapy; detection; immunogenic; cytostatic; vaccine.

Homo sapiens.

WO200004149-A2.

27-JAN-2000.

14-JUL-1999; 99WO-US15838.

14-JUL-1998; 98US-0115453.

14-JUL-1998; 98US-0116134.

23-SEP-1998; 98US-0159812.

23-SEP-1998; 98US-0159822.

15-JAN-1999; 99US-0232149.

15-JAN-1999; 99US-0232880.

09-APR-1999; 99US-0288946.

(CORI-) CORIXA CORP.

Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;

WPI; 2000-171268/15.

New polypeptide useful for treating and diagnosing prostate cancer comprises an immunogenic portion of prostate tumor protein -

Claim 3; Page 139-140; 263pp; English.

The present invention describes isolated polypeptides, comprising an immunogenic portion of a prostate tumour protein (PTP). The polypeptides and polynucleotides encoding them have cytostatic activity and can be used in vaccines and in gene therapy. The polypeptides and polynucleotides encoding them, antigen presenting cells which express the polypeptides, antibodies against the polypeptides and vaccines comprising them can be used for inhibiting the development of prostate cancer in a patient. The polypeptides can be used to generate antibodies or anti-idiotypic antibodies for passive immuno therapy. A portion of the polynucleotides encoding the polypeptides can be used as a probe or to modulate the expression of the polypeptides. AAA06241 to AAA06691 and AAV82000 to AAV82020 represent sequences used in the exemplification of the present invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 21; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

61 IAAGVVVFALGFLGCGVAKTESKCALVTFPFFILLIFIAEVAARAAVVALVTTMAEHLTL 120

61 IAAGVVVFALGFLGCGVAKTESKCALVTFPFFILLIFIAEVAARAAVVALVTTMAEHLTL 120

121 LVVPAIKDYGSGEDFTQVWNTMTKGLKCCGFTNYTDFEDSPFKENSAPPPFCCNDNVT 180

121 LVVPAIKDYGSGEDFTQVWNTMTKGLKCCGFTNYTDFEDSPFKENSAPPPFCCNDNVT 180

181 NTANETCTQKARDQVECFNQLLYDTRTNAVTVGGVAAGIGGLELAAMTVSYLYCNL 240

181 NTANETCTQKARDQVECFNQLLYDTRTNAVTVGGVAAGIGGLELAAMTVSYLYCNL 240

QY 241 Q 241

Db 241 Q 241

RESULT 6

AAU69764

ID AAU69764 standard; Protein; 241 AA.

XX AAU69764;

XX 30-JAN-2002 (first entry)

XX Human prostate cDNA encoded protein #4.

XX Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.

XX Homo sapiens.

XX WO200173032-A2.

XX 04-OCT-2001.

XX 27-MAR-2001; 2001WO-US09919.

XX 27-MAR-2000; 2000US-0536857.

XX 09-MAY-2000; 2000US-0568100.

XX 12-MAY-2000; 2000US-0570737.

XX 13-JUN-2000; 2000US-0593793.

XX 27-JUN-2000; 2000US-0605783.

XX 10-AUG-2000; 2000US-0636215.

XX 29-AUG-2000; 2000US-0651236.

XX 06-SEP-2000; 2000US-0657279.

XX 02-OCT-2000; 2000US-0679426.

XX 10-OCT-2000; 2000US-0685166.

XX (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedwick TS, Carter D;

PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

XX WPI; 2001-639232/73.

XX N-PSDB; AAS63558.

XX New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer -

PS Claim 2; Page 270-271; 579pp; English.

XX The invention relates to isolated prostate-specific

CC polynucleotides, polypeptides, fusion proteins of the polypeptides,

CC antibodies raised against the polypeptides (or antigenic epitopes

CC derived from them) and antigen-presenting cells expressing the

CC polypeptides. The antibodies are useful for detecting the presence of

CC cancer, especially prostate cancer. The polypeptides, polynucleotides and

CC the antigen-presenting cells are useful for stimulating and/or expanding

CC T cells specific for a tumour protein, and for inhibiting the development

CC of cancer especially prostate cancer. Compositions comprising the

CC polynucleotide and/or polypeptide are useful for stimulating an immune

CC response, and for treating cancer. The oligonucleotide is useful for

CC detecting cancer. The present sequence is a prostate specific

CC polypeptide of the invention.

XX Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQCFSPKTMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGLPSSAMQFVNVGYFL 60

1 MQCFSTKTMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60  
61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120  
61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120  
121 LVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPFCNDNVT 180  
121 LVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPFCNDNVT 180  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
241 Q 241  
241 Q 241

ULT 7  
04962

AAU04962 standard; Protein; 241 AA.

AAU04962;

24-OCT-2001 (first entry)

Human prostate tumour protein N1-1862.

Human; prostate tumour protein; prostate cancer.

Homo sapiens.

USG262245-B1.

17-JUL-2001.

25-FEB-1998; 98US-0030607.

25-FEB-1997; 97US-0806099.

01-AUG-1997; 97US-0904804.

09-FEB-1998; 98US-0020956.

(CORI-) CORIXA CORP.

Xu J, Dillon DC;

WPI; 2001-440862/47.

N-PSDB; AAS10109.

Novel polynucleotide encoding polypeptide comprising a portion of prostate tumour protein useful for inhibiting development of prostate cancer or for treating prostate cancer in a patient -

Example 1; Column 128-129; 105pp; English.

The sequence is a partial prostate tumour protein, encoded by a prostate tumour specific cDNA. The DNA is useful for inhibiting the development of prostate cancer or for treating prostate cancer in a patient.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132; Indels 0; Gaps 0;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MQCFSTKTMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60

1 MQCFSTKTMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60

61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120

61 IAAGVVVFALGFLGCGYKATESKCALVTFFFIILLIFIAEVAANVALVYTTWASHFLTL 120

Qy 121 LVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPFCNDNVT 180  
Db 121 LVPAIKDYGSQEDFTQVWNTTMKGLKCCGFTNYTDFEDSPYFKENSAPFPFCNDNVT 180  
Qy 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
Db 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNATVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
Qy 241 Q 241  
Db 241 Q 241

RESULT 8

AAU01118

ID AAU01118 standard; Protein; 241 AA.

XX AC AAU01118;

DT 04-OCT-2001 (first entry)

XX Human prostate-specific amino acid sequence N1-1862.

XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;  
XX KW cytostatic; gene therapy; metastasis.

XX OS Homo sapiens.

XX PN WO200151633-A2.

XX PD 19-JUL-2001.

XX PF 16-JAN-2001; 2001WO-US01574.

XX PR 14-JAN-2000; 2000US-0483672.

XX PA (CORI-) CORIXA CORP.

XX PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

XX PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;

XX PI Wang A, Meagher MJ;

XX DR WPI; 2001-425873/45.

XX PT New polynucleotide encoding a prostate-specific protein, for  
PT diagnosing, monitoring and treating prostate cancer in a patient and  
PT for use in vaccines -

XX PS Claim 2; Page 268-269; 543pp; English.

XX The present invention describes polynucleotide sequences (I) which encode  
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,  
CC and can be used in vaccine production and gene therapy. (I), (II),  
CC antibodies to (II), fusion proteins comprising (II), and isolated  
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.  
CC (I) and the antibodies are also used in the detection of cancer in a  
CC patient. The cancer that is diagnosed or treated is particularly  
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or  
CC (I) can be used for monitoring the progression of cancer in a patient.  
CC (I) and (II) can also be used to improve diagnostic and therapeutic  
CC methods for prostate cancer. They can indicate the level of metastasis  
CC as well as the prostate volume. AA93357 to AA93944 and AAU01115 to  
CC AAU01118 represent polynucleotide and amino acid sequences used in the  
CC exemplification of the present invention.

XX SQ Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;

Best Local Similarity 100.0%; Pred. No. 6.4e-132; Indels 0; Gaps 0;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQCFSTKTMILFNLLIFLCGALLAVGIWVSIDGASFLKIFGPISSSAMQFVNVGYFL 60



|||||  
1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPIUSSSAMQFVNVGYFL 60  
61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
241 Q 241  
241 Q 241

ULT 9  
399003  
AAG99003 standard; Protein; 241 AA.

AAG99003;

25-SEP-2001 (first entry)

Human prostate-specific amino acid sequence N1-1862/P503S.

Human; prostate cancer; therapy; diagnosis; cat eye syndrome;  
chromosome 22q11.2; prostate-specific protein; chromosome 1;  
prostate specific antigen; PSA.

Homo sapiens.

WO200134802-A2.

17-MAY-2001.

09-NOV-2000; 2000WO-US30904.

12-NOV-1999; 99US-0439313.

18-NOV-1999; 99US-0443686.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;  
Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;

WPI; 2001-308785/32.

Isolated polypeptide comprising at least an immunogenic portion of a  
prostate-specific protein, useful in the diagnosis and therapy of  
prostate cancer -

Claim 3; Page 168-169; 325pp; English.

The present invention describes an isolated polypeptide (P1) comprising  
at least an immunogenic portion of a prostate-specific protein, or its  
variant. Also described are polynucleotides (N1) encoding (P1). (P1) and  
(N1) have cytostatic activity and can be used in vaccine production.  
The polypeptides, nucleic acids and antibodies from the present  
invention are useful in the diagnosis and therapy of prostate cancer.  
Prostate specific genes P704P, P712P, P774P, P775P and B305D are located  
in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome  
region. Prostate specific antigen (PSA) P501S was located on  
chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent  
polynucleotide and polypeptide sequences used in the exemplification  
of the present invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132; Indels 0; Gaps 0;  
Matches 241; Conservative 0; Mismatches 0;  
QY 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPIUSSSAMQFVNVGYFL 60  
DB 1 MQCFSTKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGPIUSSSAMQFVNVGYFL 60  
QY 61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
DB 61 IAAGVVFALGFLGCGYGAKTESKCALVTFFILLIFIAEVAALVALVYTTMAEHFLTL 120  
QY 121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
DB 121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
QY 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
DB 181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
QY 241 Q 241  
DB 241 Q 241

RESULT 10

AAAB74801

ID AAB74801 standard; Protein; 241 AA.

AC AAB74801;

XX 14-JUN-2001 (first entry)

Prostate tumour antigen predicted amino acid sequence for N1-1862.

Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;  
Prostate cancer; immunogenic; cytostatic; vaccine.

OS Homo sapiens.

PN WO200125272-A2.

PD 12-APR-2001.

PF 04-OCT-2000; 2000WO-US27464.

PR 04-OCT-1999; 99US-0157455.

PA (CORI-) CORIXA CORP.

PI Xu J, Skeiky YAW, Reed SG, Cheever MA;

XX WPI; 2001-245062/25.

DR N-PSDB; AAH02531.

PT Prostate specific protein and its encoding polynucleotide, useful for  
the treatment and diagnosis of prostate cancer -

XX Claim 3; Page 158-159; 276pp; English.

PS The present invention describes an isolated polypeptide (I) comprising  
at least an immunogenic portion of a prostate tumour antigen protein or  
its variant. (I) have cytostatic activity and can be used in vaccine  
production. (I), prostate tumour antigen polynucleotides, an antigen  
presenting cell (APC e.g. a dendritic cell) that expresses (I), and a  
pharmaceutical composition containing (I) are useful for inhibiting the  
development of cancer in a patient. Antibodies specific for prostate  
specific proteins and oligonucleotides that hybridise to a  
polynucleotide that encodes a prostate specific protein are useful  
for detecting the presence or absence of a cancer or monitoring the  
progression the progression of a cancer, especially prostate cancer.  
AAH02422 to AAH2872; AAB74798 to AAB74821 and AAB74830 are sequences  
used in the exemplification of the present invention.

Sequence 241 AA;  
 Query Match 100.0%; Score 1258; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MQCFSEIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGFLSSAMQFVNVGYFL 60  
 1 MQCFSEIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGFLSSAMQFVNVGYFL 60  
 61 IAGVVVFALGFLGCGYKATESKCALVTFFPILLIFIAEVAAVVALVYTTMAEHFLL 120  
 61 IAGVVVFALGFLGCGYKATESKCALVTFFPILLIFIAEVAAVVALVYTTMAEHFLL 120  
 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDVT 180  
 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDVT 180  
 181 NTANETCTKQAHDKQVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQAHDKQVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

UT 11  
 90679

AAB90679 standard; Protein; 241 AA.

AAB90679;

07-JUN-2001 (first entry)

Human AR415\_4 protein sequence SEQ ID 35.

Human; secreted protein; nutrient; cytokine modulator; proliferation;  
 differentiation; immune system modulator; tissue growth; chemotactic;  
 haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 haematopoiesis.

Homo sapiens.

WO200119388-A1.

22-MAR-2001.

14-SEP-2000; 2000WO-US25135.

17-SEP-1999; 99US-0398829.

(GEMY) GENETICS INST INC.

Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C,  
 Merberg D, Treacy M, Bowman MR, Spaulding V, Agostino MJ;

WPI: 2001-244801/25.

N-PSDB; AAF98395.

Isolated nucleic acids encoding polypeptides, useful for modulating  
 e.g. cytokine and cell proliferation/differentiation activity, the  
 immune system and hematopoiesis regulating activity -

Disclosure; Page 403-404; 557pp; English.

Human cDNA clones represented in AAF98374 - AAF98489 encode secreted  
 proteins AAB90667 - AAB90750. The cDNA clones are isolated from various  
 tissue types, and may be used in the prevention, treatment and diagnosis  
 of diseases associated with inappropriate protein expression. The  
 polypeptides and nucleic acids may be used as nutrients or to modulate  
 cytokine and cell proliferation/differentiation activity and may also be

CC involved in modulation of the immune system. The cDNA sequences,  
 proteins, their agonists and/or antagonists exhibit haematopoiesis,  
 regulating activity; tissue growth activity; activin/inhibin activity;  
 chemotactic/chemokinetic activity; haemostatic and thrombolytic  
 activity; receptor/ligand activity; anti-inflammatory activity;  
 haematopoiesis activity; cadherin/tumour suppressor activity; and/or  
 tumour inhibition activity. Included in the invention are probes  
 CC represented in AAF98490 - AAF98572 which are specific for the cDNA clones  
 CC encoding the secreted proteins.

SQ Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 22; Length 241;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQCFSEIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGFLSSAMQFVNVGYFL 60

DB 1 MQCFSEIKTMMILFNLLIFLCGAALLAVGIWISIDGASFLKIFGFLSSAMQFVNVGYFL 60

QY 61 IAGVVVFALGFLGCGYKATESKCALVTFFPILLIFIAEVAAVVALVYTTMAEHFLL 120

DB 61 IAGVVVFALGFLGCGYKATESKCALVTFFPILLIFIAEVAAVVALVYTTMAEHFLL 120

QY 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDVT 180

DB 121 LVVPAIKDYGSQEDFTQVWNTTMMGLKCCGFTNYTDFEDSPYFKENSAPFPCCNDVT 180

QY 181 NTANETCTKQAHDKQVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240

DB 181 NTANETCTKQAHDKQVEGCFNQLLYDIRTNVTVGGVAAGIGGLELAAMIVSMYLYCNL 240

QY 241 Q 241

DB 241 Q 241

RESULT 12

ABG94412

ID ABG94412 standard; Protein; 241 AA.

XX ABG94412;

DT 27-NOV-2002 (first entry)

DE Human prostate tumour protein partial sequence #4.

XX Human; immunogenic; prostate protein; prostate tumour protein;

KW prostate cancer; cytostatic; vaccine.

XX Homo sapiens.

XX US2002030372-A1.

XX 11-JUL-2002.

XX 14-JUL-1998; 98US-0115453.

XX 25-FEB-1997; 97US-0806099.

PR 01-AUG-1997; 97US-0904804.

PR 03-FEB-1998; 98US-0020956.

PR 25-FEB-1998; 98US-0030607.

XX (XUJJ/) XU J.

PA (DILL/) DILLON D C.

XX Xu J, Dillon DC;

XX WPI; 2002-642373/69.

DR N-PSDB; ABS21255.

XX Novel polypeptides useful as vaccines for inhibiting prostate cancer  
 PT development, comprise an immunogenic portion of prostate protein -

Example 1; Page 61-62; 101pp; English.

The present invention relates to a new polypeptide comprising an immunogenic portion of a prostate protein. The invention is useful for inhibiting the development of prostate cancer in a patient. The invention is also useful as markers for diagnosing prostate cancer and for monitoring diseases progression in patients. The present amino acid sequence represents a human prostate tumour protein.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 23; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MQCFSFIKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60  
1 MQCFSFIKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60  
61 IAAGVWVVFALGFLGCGYKAKTESKCALVTFFILLIFIAEVAADVVALVTTWAEHFLTL 120  
61 IAAGVWVVFALGFLGCGYKAKTESKCALVTFFILLIFIAEVAADVVALVTTWAEHFLTL 120  
121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
241 Q 241  
241 Q 241

SULT 13  
376666  
ABG76666 standard; Protein; 241 AA.  
ABG76666;

05-NOV-2002 (first entry)  
Prostate tumour protein #4.  
Human; prostate tumour; immunotherapy; prostate cancer.  
Homo sapiens.

US2002081580-A1.

27-JUN-2002.

25-FEB-1998; 98US-0030606.

25-FEB-1997; 97US-0806596.

01-AUG-1997; 97US-0904809.

09-FEB-1998; 98US-0020747.

(XUJG/) XU J.

(DILL/) DILLON D C.

Xu J, Dillon DC;

WPI; 2002-607662/65.

Detecting prostate cancer comprises contacting a sample with an agent capable of binding to a polypeptide with an immunogenic portion of a prostate protein, oligonucleotide primers or a probe specific for DNA encoding the polypeptide

Example 1; Page 67-68; 111pp; English.

The invention relates to a method of detecting prostate cancer by contacting a biological sample from a patient with: (a) a binding agent that binds to a polypeptide having an immunogenic portion of a prostate protein or its variant; (b) 2 oligonucleotide primers, where 1 of the oligonucleotides is specific for a DNA encoding the polypeptide of (a); or (c) an oligonucleotide probe specific for a DNA molecule encoding the polypeptide of (a). The method and polypeptides are useful for diagnosing, treating, particularly by immunotherapy, monitoring the progression, and inhibiting the development of prostate cancer in a patient. The polypeptides may be used to generate antibodies useful for the diagnosis and monitoring of prostate cancer. ABG76663-ABG76669 represent human prostate tumour protein sequences of the invention.

Sequence 241 AA;

Query Match 100.0%; Score 1258; DB 23; Length 241;  
Best Local Similarity 100.0%; Pred. No. 6.4e-132;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 MQCFSFIKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60  
1 MQCFSFIKTMILFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLPLSSAMQFVNVGYFL 60  
61 IAAGVWVVFALGFLGCGYKAKTESKCALVTFFILLIFIAEVAADVVALVTTWAEHFLTL 120  
61 IAAGVWVVFALGFLGCGYKAKTESKCALVTFFILLIFIAEVAADVVALVTTWAEHFLTL 120  
121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
121 LVVPAIKDYGSQEDFTQVWNTTKGLKCCGFTNYTDFEDSPYFKENSAPPPCCNDNVT 180  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGGVAAGIGLELAAMIVSMYLYCNL 240  
241 Q 241  
241 Q 241

RESULT 14  
ABB95223  
ID ABB95223 standard; Protein; 241 AA.  
AC ABB95223;

19-JUL-2002 (first entry)

Human N1-1862 protein SEQ ID NO 114.

Human; cancer; prostate cancer; vaccine; cytostatic; immunostimulant; gene therapy.

Homo sapiens.

US2002022248-A1.

21-FEB-2002.

12-JAN-2001; 2001US-0759143.

25-FEB-1997; 97US-0806099.

01-AUG-1997; 97US-0904804.

09-FEB-1998; 98US-0020956.

25-FEB-1998; 98US-0030607.

14-JUL-1998; 98US-0115453.

23-SEP-1998; 98US-0159812.

15-JAN-1999; 99US-0232149.

09-APR-1999; 99US-028946.

13-JUL-1999; 99US-032616.

12-NOV-1999; 99US-0439313.

18-NOV-1999; 99US-0443686.  
 14-JAN-2000; 2000US-0483672.  
 27-MAR-2000; 2000US-0536857.  
 09-MAY-2000; 2000US-0568100.  
 12-MAY-2000; 2000US-0570737.  
 13-JUN-2000; 2000US-0593793.  
 27-JUN-2000; 2000US-0605783.  
 10-AUG-2000; 2000US-0636215.  
 29-AUG-2000; 2000US-0651236.  
 06-SEP-2000; 2000US-0657279.  
 02-OCT-2000; 2000US-0679426.  
 10-OCT-2000; 2000US-0685166.

(XUJJ/) XU J.  
 (DILL/) DILLON D C.  
 (MITC/) MITCHAM J L.  
 (HARL/) HARLOCKER S L.  
 (JIAN/) JIANG Y.  
 (KALO/) KALOS M D.  
 (FANG/) FANGER G R.  
 (RETT/) RETTER M W.  
 (STOL/) STOLK J A.  
 (DAYC/) DAY C H.  
 (VEDV/) VEDVICK T S.  
 (CART/) CARTER D.  
 (LISX/) LI S X.  
 (WANG/) WANG A.  
 (SKEI/) SKEIKY Y A W.  
 (HEPL/) HEPLER W T.  
 (HEND/) HENDERSON R A.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

WPI; 2002-255649/30.

New prostate-specific polynucleotides for diagnosing and treating  
 diseases, in particular prostate cancer, and as markers for the  
 progression of cancer

Claim 2; SEQ ID NO 114; 87pp; English.

The present invention provides prostate-specific coding sequences and  
 their encoded proteins. These can be used in the diagnosis and treatment  
 of cancers, particularly prostate cancer. The present sequence is a  
 protein described in the invention.

Sequence 241 AA;

Very Match 100.0%; Score 1258; DB 23; Length 241;  
 est Local Similarity 100.0%; Pred. No. 6.4e-132;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MOCSEFTKMWLENAFLCGALLAVGTWSDGASFLKIFGPISSAMQFVNVGYFL 60  
 1 MOCSEFTKMWLENAFLCGALLAVGTWSDGASFLKIFGPISSAMQFVNVGYFL 60  
 61 IAAGVWVFALGFLGCGAKTESKCALVTPFFILLIFIAEVAAVVALVYTTMAEHFTL 120  
 61 IAAGVWVFALGFLGCGAKTESKCALVTPFFILLIFIAEVAAVVALVYTTMAEHFTL 120  
 121 LVWPAIKDYGSDFTQVNTWTKGLKCGFNFTDFEDSPYFKENSAPFFCCNDNVT 180  
 121 LVWPAIKDYGSDFTQVNTWTKGLKCGFNFTDFEDSPYFKENSAPFFCCNDNVT 180  
 181 NTANETCTKQAHQKQEGGCFNQLLYDIRNAVTVGVGAAGIGGLELAAMIVSMYLYCNL 240  
 181 NTANETCTKQAHQKQEGGCFNQLLYDIRNAVTVGVGAAGIGGLELAAMIVSMYLYCNL 240  
 241 Q 241  
 241 Q 241

# RESULT 15

ABU71654 standard; Protein; 241 AA.

XX AC ABU71654;  
 XX DT 10-JUN-2003 (first entry)  
 XX DE Prostate cancer associated protein #3.  
 XX KW Prostate cancer; vaccine; gene therapy; cytostatic; fusion protein;  
 KW immunogen; cancer; prostate specific antigen; PSA; membrane antigen;  
 KW prostatic acid phosphatase; PAP; prostate specific membrane antigen;  
 PSMA.

OS Homo sapiens.

EN US2002192763-A1.

XX PD 19-DEC-2002.

XX PF 29-JUN-2001; 2001US-0895793.

XX PR 17-APR-2000; 2000US-157455P.

XX PR 04-OCT-2000; 2000US-0679272.

XX PR 28-MAR-2001; 2001US-0822827.

XX PA (XUJJ/) XU J.

XX PA (DILL/) DILLON D C.

XX PA (MITC/) MITCHAM J L.

XX PA (HARL/) HARLOCKER S L.

XX PA (JIAN/) JIANG Y.

XX PA (KALO/) KALOS M D.

XX PA (FANG/) FANGER G R.

XX PA (RETT/) RETTER M W.

XX PA (STOL/) STOLK J A.

XX PA (DAYC/) DAY C H.

XX PA (VEDV/) VEDVICK T S.

XX PA (CART/) CARTER D.

XX PA (LISX/) LI S X.

XX PA (WANG/) WANG A.

XX PA (SKEI/) SKEIKY Y A W.

XX PA (HEPL/) HEPLER W T.

XX PA (HEND/) HENDERSON R A.

XX PA (HURA/) HURAL J.

XX PA (MCNE/) MCNEILL P D.

XX PA (HOUG/) HOUGHTON R L.

XX PA (DBAS/) Y DE BASSOLS C V.

XX PA (FOYT/) FOY T M.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;  
 Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;  
 Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA, Hural J;  
 McNeill PD, Houghton RL, Y De Bassols CV, Foy TM;  
 WPI; 2003-352711/33.

XX New fusion protein comprising prostate-specific polypeptides, or its  
 immunogenic portions, useful for diagnosing, preventing and/or treating  
 cancer, particularly prostate cancer

Example 1; SEQ ID NO 114; 85pp; English.

The invention describes a fusion protein comprising at least one amino  
 acid sequence of immunogenic portions of any of the 3 sequences not  
 defined in the specification, or sequences having at least 70 or 90 %  
 sequence identity to any one of the 35 sequences defined in the USPTO  
 web site, which is encoded by any of the 4 nucleotide sequences not  
 defined in the specification. The fusion protein, composition and  
 methods are useful for diagnosing, preventing and/or treating cancer,  
 particularly prostate cancer. The proteins are useful as markers to

indicate the presence or absence of cancer. This is the amino acid sequence of a prostate cancer therapy associated protein.  
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US patent office at  
seqdata.uspto.gov/sequence.html?DocID=US20020192763.

```
Sequence 241 AA;
Query Match 100.0%; Score 1258; DB 24; Length 241;
Est Local Similarity 100.0%; Pred. No. 6.4e-132;
atches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MQCPSFIKTWMLFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLSSAMQFVNVGYFL 60
1 MQCPSFIKTWMLFNLLIFLCGAALLAVGIWVSIDGASFLKIFGLSSAMQFVNVGYFL 60
61 IAAGVWVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAAVVALVYTTMAEHFLTL 120
61 IAAGVWVFALGFLGCGYKATESKCALVTFPFIILLIFIAEVAAVVALVYTTMAEHFLTL 120
121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDPEDSPYFKENSAPFPCCNDNVT 180
121 LVVPAIKDYGSQEDFTQVWNTTMMKGLKCCGFTNYTDPEDSPYFKENSAPFPCCNDNVT 180
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGVGAAGIGGLELAAMIVSMYLYCNL 240
181 NTANETCTKQKAHDQKVEGCFNQLLYDIRTNVTVGVGAAGIGGLELAAMIVSMYLYCNL 240
241 Q 241
241 Q 241
```

rch completed: February 2, 2004, 09:18:03  
ime : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 4, 2004, 14:09:14 ; Search time 47 Seconds

(without alignments)  
3789.175 Million cell updates/sec

le: US-09-042-460-2

fect score: 5901

uence: 1 NTRAPRCPAVSLRSRYRE.....TILKARADPALSTDFQTILD 1122

ring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

urched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

imum DB seq length: 0

cimum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

A\_Geneseq 19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

alt No.	Score	Match	Length	ID	Description
1	5901	100.0	1122	20 AAY26579	Murine telomerase
2	5854	99.2	1122	23 ABB06711	Mouse telomerase p
3	3505	59.4	1132	19 AAW46957	Human telomerase PA
4	3505	59.4	1132	20 AAY43621	A human telomerase
5	3505	59.4	1132	20 AAY28881	Human telomerase r
6	3505	59.4	1132	20 AAY32090	Human telomerase r
7	3505	59.4	1132	20 AAY26580	Human telomerase r
8	3505	59.4	1132	20 AAW90251	Human catalytic te
9	3505	59.4	1132	22 AAB82765	Human telomerase r

10	3505	59.4	1132	22 AAB99930	Human telomerase p
11	3505	59.4	1132	22 AAG64329	Human protein #2.
12	3505	59.4	1132	22 AAG64859	Heart muscle cell
13	3505	59.4	1132	23 AAR29226	Human telomerase r
14	3505	59.4	1132	23 ABU72735	Human telomerase r
15	3505	59.4	1132	24 ABP56676	Human telomerase r
16	3505	59.4	1154	19 AAW61350	Human telomerase p
17	3505	59.4	1189	19 AAW47008	Glutathione-S-tran
18	3499	59.3	1285	19 AAW47000	HIS tagged thior
19	3498	59.3	1132	19 AAW71376	Human telomerase c
20	3498	59.3	1132	20 AAY28401	Human ESR2 protein
21	3498	59.3	1132	20 AAY00627	Human telomerase p
22	3498	59.3	1132	20 AAY00638	Truncated telomera
23	3498	59.3	1132	21 AAY96566	hEST2, a human tel
24	3496	59.2	1132	19 AAW56113	Human telomerase r
25	3486.5	59.1	1199	19 AAW47007	Glutathione-S-tran
26	3471	58.8	1166	20 AAY00647	Enhanced green flu
27	3463	58.7	1405	19 AAW56101	Telomerase (ver. 2
28	3435	58.2	1120	20 AAY00641	Telomerase (ver. 2
29	3426	58.1	1120	20 AAY00650	Telomerase (ver. 2
30	3320	56.3	1150	19 AAW47006	Glutathione-S-tran
31	3277.5	55.5	1053	20 AAY00640	Altered C-terminus
32	3239.5	54.9	1093	20 AAY00649	Altered C-terminus
33	3198.5	54.2	1041	20 AAY00643	Altered C-terminus
34	3198.5	54.2	1041	20 AAY00652	Altered C-terminus
35	2881.5	48.8	948	20 AAY00639	N-terminal truncat
36	2877.5	48.8	948	20 AAY00648	Truncated telomera
37	2814.5	47.7	936	20 AAY00642	Truncated telomera
38	2805.5	47.5	936	20 AAY00651	Truncated telomera
39	2790.5	47.3	949	19 AAW61349	Human telomerase p
40	2195	37.2	807	19 AAW46997	Human telomerase r
41	2195	37.2	807	20 AAY00637	N-terminal truncat
42	2195	37.2	807	20 AAY00646	Truncated telomera
43	2132	36.1	622	20 AAY25463	Human CRT-1 protei
44	2073	35.1	591	20 AAW97384	A catalytic telome
45	1709	29.0	499	24 AAB999678	Amino acid sequenc

#### ALIGNMENTS

RESULT 1  
AAY26579  
ID AAY26579 standard; Protein; 1122 AA.  
AC  
AAY26579;  
AC  
DT 13-SEP-1999 (first entry)  
XX  
DE Murine telomerase reverse transcriptase (mTERT) enzyme.  
XX  
XX Telomerase reverse transcriptase; TERT; mouse; telomere length assay;  
KW immunogen; enzyme; telomerase-mediated DNA replication.  
XX  
OS Mus sp.  
XX  
XX WO9927113-A1.  
XX  
XX 03-JUN-1999.  
XX  
XX 25-NOV-1998; 98WO-US25211.  
XX  
XX 16-MAR-1998; 98US-0042460.  
XX 26-NOV-1997; 97US-0979742.  
XX  
XX (GERO-) GERON CORP.  
XX (YESH) UNIV YESHIVA EINSTEIN COLLEGE.  
XX  
XX Allisopp R, Depinho R, Greenberg R, Morin GB;  
XX WPI, 1999-347722/29.  
XX N-PSDB; AAX80994.  
XX

Mouse telomerase reverse transcriptase (mTERT) enzyme proteins and nucleic acids

Claim 8; Fig 2; 135pp; English.

The invention relates to a mouse telomerase reverse transcriptase (mTERT) enzyme. Compositions containing mTERT can be used in telomere length assays. Isolated mTERT is useful as an immunogen for the production of monoclonal or polyclonal antibodies. The method is useful for assessing the degree of purification and identification of new mTERT species, such as an mTERT allele, homolog or isoform, or to screen for modulators (antagonists and agonists) of telomerase-mediated DNA replication. Other telomerase enzymes such as human TERT (hTERT). The present sequence represents a mTERT enzyme.

Sequence 1122 AA;

Query Match 100.0%; Score 5901; DB 20; Length 1122;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MTRAPCPAVRSLLRSRYREVWPLATFVRLOPEGRRLVQGPDKIYRTLVAQCLVCHW 60  
1 MTRAPCPAVRSLLRSRYREVWPLATFVRLOPEGRRLVQGPDKIYRTLVAQCLVCHW 60  
61 GSOPPPADLSFHQVSSIKELVAVQVORLCERNVLAFFELLNARGGPPMAFTSSVR 120  
61 GSOPPPADLSFHQVSSIKELVAVQVORLCERNVLAFFELLNARGGPPMAFTSSVR 120  
121 SYLPNTVIETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPNTVIETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASYPTRPVGNFTNLFLOIKSSQRBAPKPLALPSRGTKEHLSLTSTS 240  
181 TTDIWPVSASYPTRPVGNFTNLFLOIKSSQRBAPKPLALPSRGTKEHLSLTSTS 240  
241 VPSAKKARCPVPRVEGPHRQVLPFPSSGKSWPSPARSPVEPTAEKOLSSKGVSDLSL 300  
241 VPSAKKARCPVPRVEGPHRQVLPFPSSGKSWPSPARSPVEPTAEKOLSSKGVSDLSL 300  
301 SGVCCCKHPSSTLSLSPQNAFQRLPFIETHFLYSRGDQERLNPSFLLSNLQPNLT 360  
301 SGVCCCKHPSSTLSLSPQNAFQRLPFIETHFLYSRGDQERLNPSFLLSNLQPNLT 360  
361 GARLVEIIFLGSRPSTGPLCTHRLSRYWOMRPLFOQLLVNHAECQYVLLRSHCRF 420  
361 GARLVEIIFLGSRPSTGPLCTHRLSRYWOMRPLFOQLLVNHAECQYVLLRSHCRF 420  
421 RTANQOQVTDALNTSPPHMDLLRLHSPWQVYGLRACLCVKVVSASLWTRNRRFPKN 480  
421 RTANQOQVTDALNTSPPHMDLLRLHSPWQVYGLRACLCVKVVSASLWTRNRRFPKN 480  
481 LKKTISLCKYKLSLQELMKMKVDCWHLRSSPGKDRVPAEHLRLERILATFLWMD 540  
481 LKKTISLCKYKLSLQELMKMKVDCWHLRSSPGKDRVPAEHLRLERILATFLWMD 540  
541 TVVQQLRSFFYITESTFQKNLFFYRKSWLSQIGVROHLERVRLRELSQEBVRHQ 600  
541 TVVQQLRSFFYITESTFQKNLFFYRKSWLSQIGVROHLERVRLRELSQEBVRHQ 600  
601 DTWLAMPICRLRFTPKNGRLPIVNSYMGTRALGRKQAOHFTORLKLTLFSLMYERT 660  
601 DTWLAMPICRLRFTPKNGRLPIVNSYMGTRALGRKQAOHFTORLKLTLFSLMYERT 660  
661 KPHLMGSSVLGMDIIVTRWAFVLRVLRALDQTPRMVFKADVTGAYDAIPQKLVVEVA 720  
661 KPHLMGSSVLGMDIIVTRWAFVLRVLRALDQTPRMVFKADVTGAYDAIPQKLVVEVA 720  
721 NWIRHSESTYCIRQYAVVRDSDQGVHKSFRQVTTLSLQIPMGQFLKHLQSDASALR 780  
721 NWIRHSESTYCIRQYAVVRDSDQGVHKSFRQVTTLSLQIPMGQFLKHLQSDASALR 780

QY 781 NSVIEQISINWESSSLFDFFLHFLRHSVVKIGDRCYTCQCGIPQSSSLTLLCSLCFG 840  
DB 781 NSVIEQISINWESSSLFDFFLHFLRHSVVKIGDRCYTCQCGIPQSSSLTLLCSLCFG 840  
QY 841 DMENKLFPAEVQDGLLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGCINLQKTVVNF 900  
DB 841 DMENKLFPAEVQDGLLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGCINLQKTVVNF 900  
QY 901 PVEPOTLGAAPYQLPDAHCLFPWCGLLLDQTLEVFCDYSGYACTSIKTLTFSQVFKAG 960  
DB 901 PVEPOTLGAAPYQLPDAHCLFPWCGLLLDQTLEVFCDYSGYACTSIKTLTFSQVFKAG 960  
QY 961 KTMRNKLLSVLRKCHGLFLLDQVNSLQTVCIINIKIFLLQAYRFAHCVIQLPFDQVRK 1020  
DB 961 KTMRNKLLSVLRKCHGLFLLDQVNSLQTVCIINIKIFLLQAYRFAHCVIQLPFDQVRK 1020  
QY 1021 NLTFGLGISSQASCCYAILKVNPGMTLKASGFPPEAAHWLCYQAFLLKLAHSHVIYK 1080  
DB 1021 NLTFGLGISSQASCCYAILKVNPGMTLKASGFPPEAAHWLCYQAFLLKLAHSHVIYK 1080  
QY 1081 CULGELRTAQKLLCKKPEATMTILKAAADPALSTDFQTILD 1122  
DB 1081 CULGELRTAQKLLCKKPEATMTILKAAADPALSTDFQTILD 1122

RESULT 2

ABB06711

ID ABB06711 standard; Protein; 1122 AA.

XX ABB06711;

XX 11-JUN-2002 (first entry)

XX Mouse telomerase protein sequence.

XX Mouse; telomerase; promoter; telomerase catalyst subunit; TERT; mTERT;

XX enzyme; transgenic mouse; drug development; anticancer.

XX Mus sp.

XX JP2002000121-A.

XX 08-JAN-2002.

XX 23-JUN-2000; 2000JP-0190137.

XX 23-JUN-2000; 2000JP-0190137.

XX (RIKO-) ZH RIKOGAKU SHINKOKAI.

XX (KIRI) KIRIN BREWERY KK.

XX WPI; 2002-298279/34.

XX A transgenic mouse comprising a DNA promoter region of mouse telomerase catalyst subunit (TERT) is used for the development of drugs and anticancer agents for regeneration of tissues and organs -

XX Disclosure; Fig 3; 13pp; Japanese.

XX The present invention describes a transgenic mouse (I) comprising a DNA construct having a DNA containing a promoter region of mouse telomerase catalyst subunit (TERT) and a DNA containing a reporter gene connected under the control of the promoter region. The transgenic mouse can be used in the development of drugs and anticancer agents for regeneration of tissues and organs. The present sequence represents the mouse telomerase protein, which is given in the exemplification of the present invention.

XX Sequence 1122 AA;

Query Match 99.2%; Score 5854; DB 23; Length 1122;

Best Local Similarity 99.4%; Pred. No. 0;



atches 1115; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

1 MTRAPCPAVRSLLRSRYREVPVPLATFVRRLGEGRLVQPGDKPIYRTLVAQCIVCMHW 60  
1 MTRAPCPAVRSLLRSRYREVPVPLATFVRRLGEGRLVQPGDKPIYRTLVAQCIVCMHW 60

61 GSOPPPADLSFHOVSSLKELVARVQORLERNERNVLAFGFELLNARGGPPMAFTSSVR 120  
61 GSOPPPADLSFHOVSSLKELVARVQORLERNERNVLAFGFELLNARGGPPMAFTSSVR 120

121 SYLPNTVIETLRVSGAMWLLSRLVGGDDLLVYLAHCAVLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPNTVIETLRVSGAMWLLSRLVGGDDLLVYLAHCAVLLVPPSCAYQVCGSPLYQICA 180

181 TTDIWPVSASVAPTRPVGNFNLFLQOIKSSSRQEAQPKPLAPSPSGTKHLSLTSTS 240  
181 TTDIWPVSASVAPTRPVGNFNLFLQOIKSSSRQEAQPKPLAPSPSGTKHLSLTSTS 240

241 VPSAKKARCYFVPRVEGPHRQVLPPTSGKSWVPSPARSPVPTAEKDLSSKGKVDLSL 300  
241 VPSAKKARCYFVPRVEGPHRQVLPPTSGKSWVPSPARSPVPTAEKDLSSKGKVDLSL 300

301 SGSVCCCHKKSSTSLSPRONAFQIRPPIETRHFLYSRQDQGERLNFSLLSNLPNLT 360  
301 SGSVCCCHKKSSTSLSPRONAFQIRPPIETRHFLYSRQDQGERLNFSLLSNLPNLT 360

361 GARLVEIIFLGSRPRTSGPLCTRHSRYWQMRPLFQOLLVNHAECCOYVRLLSHCRF 420  
361 GARLVEIIFLGSRPRTSGPLCTRHSRYWQMRPLFQOLLVNHAECCOYVRLLSHCRF 420

421 RTANQOVTDALNTPPHLMOLLHLSPQVYGLFACICKVVSASLWGTNRNRRFFKN 480  
421 RTANQOVTDALNTPPHLMOLLHLSPQVYGLFACICKVVSASLWGTNRNRRFFKN 480

481 LKPFISLGKYGKLSQELMWMKVEDCHWLRSPGKDRVPAAEHRLRERILATFLFWM 540  
481 LKPFISLGKYGKLSQELMWMKVEDCHWLRSPGKDRVPAAEHRLRERILATFLFWM 540

541 TYVVOLLRSFFYITESTFQNRLLFFYKRSWMSKLSIGVROHLRVLRLSQQEVRHHQ 600  
541 TYVVOLLRSFFYITESTFQNRLLFFYKRSWMSKLSIGVROHLRVLRLSQQEVRHHQ 600

601 DTWLAMPICRLRFPKPNGLRPIVNNSSYSGTALGRKQAOHFTQRLKTLFSLMNYERT 660  
601 DTWLAMPICRLRFPKPNGLRPIVNNSSYSGTALGRKQAOHFTQRLKTLFSLMNYERT 660

661 KPHLMGSSVLGWNDDIYRTWRAFLVRLALDQTPRMVFKADVTGAYDAIPOGKLVVVA 720  
661 KPHLMGSSVLGWNDDIYRTWRAFLVRLALDQTPRMVFKADVTGAYDAIPOGKLVVVA 720

721 NMRHSESTYCIROYAVRRDSQGVHKSFRQVTTLSLQPYMGQFLKHLQSDASALR 780  
721 NMRHSESTYCIROYAVRRDSQGVHKSFRQVTTLSLQPYMGQFLKHLQSDASALR 780

781 NSVVIQSIWNESSSLFDFHFLHLSHVVKIGDCYTCQCGIPGSSLSLTLCSLFG 840  
781 NSVVIQSIWNESSSLFDFHFLHLSHVVKIGDCYTCQCGIPGSSLSLTLCSLFG 840

841 DMENKLPFAVQRODGLLRFDVDFLLVTHPDQAKTFLSTLVHGVPEYGCWMLQKTVNF 900  
841 DMENKLPFAVQRODGLLRFDVDFLLVTHPDQAKTFLSTLVHGVPEYGCWMLQKTVNF 900

901 PVEPGLTGGAPYOLPAHCLFPKCGLLDTQTLEVCDSYGAQTSIKTSLNFSQSVKAG 960  
901 PVEPGLTGGAPYOLPAHCLFPKCGLLDTQTLEVCDSYGAQTSIKTSLNFSQSVKAG 960

961 KTMNKLLSVLRKCHGLFDLQVNSLQTVICINIKIFLLQAYRPHACVQLPFDQVRK 1020  
961 KTMNKLLSVLRKCHGLFDLQVNSLQTVICINIKIFLLQAYRPHACVQLPFDQVRK 1020

1021 NLTFFLGIISSQASCCVAILKVRNPGMTLKASGSPPEAAHMLCYQAFLLKLAHSVYIK 1080  
1021 NLTFFLGIISSQASCCVAILKVRNPGMTLKASGSPPEAAHMLCYQAFLLKLAHSVYIK 1080

Qy 1081 CLIGPLRTAQLCKLCKLPEATWTILKAAADPALSTDFQTILD 1122  
Db 1081 CLIGPLRTAQLCKLCKLPEATWTILKAAADPALSTDFQTILD 1122

RESULT 3  
AAW46957  
ID AAW46957 standard; Protein; 1132 AA.  
XX AAW46957;  
XX AC  
XX DT 13-AUG-1998 (first entry)  
XX Human telomerase reverse transcriptase.  
XX Human; telomerase reverse transcriptase; hTERT; TRT; diagnosis;  
KW prognosis; cell proliferation; cancer; ageing; ribonucleoprotein.  
XX Homo sapiens.  
XX GB2317891-A.  
XX 08-APR-1998.  
XX 01-OCT-1997; 97GB-0020890.  
XX 14-AUG-1997; 97US-0915503.  
PR 01-OCT-1996; 96US-0724643.  
PR 18-APR-1997; 97US-0844419.  
PR 25-APR-1997; 97US-0846017.  
PR 06-MAY-1997; 97US-0851843.  
PR 09-MAY-1997; 97US-0854050.  
PR 14-AUG-1997; 97US-0911312.  
XX 14-AUG-1997; 97US-0912951.  
XX (GERO-) GERON CORP.  
PA (UYTE-) UNIV TECHNOLOGY CORP.  
XX Andrews WH, Cech TR, Chapman KB, Harley C, Lingner J;  
PI Morin GB, Nakamura T, Harley CB;  
FI N-PSDB; AAV22379.  
XX WPI; 1998-171633/16.  
DR N-PSDB; AAV22379.  
XX Pure and recombinant human Telomerase Reverse Transcriptase and its  
PT variants - are useful in the diagnosis, prognosis and treatment of  
PT cell proliferation conditions especially cancer and ageing  
XX Claim 3; Fig 17; 387pp; English.  
XX The present sequence represents human telomerase reverse transcriptase  
CC (hTERT), which is a ribonucleoprotein. The present invention also  
CC describes the following methods: (A) determining whether a test compound  
CC is a modulator of hTERT, by detecting the change in hTERT recombinant  
CC protein or polynucleotide, on administration of the compound; (B)  
CC preparation of recombinant telomerase by contacting a protein  
CC preparation of hTERT with a telomerase RNA component; (C) detection of  
CC the hTERT RNA or protein in a sample by binding a relevant probe to the  
CC sample and detecting the complex formed or in the case of RNA detection,  
CC amplifying the product and correlating the presence of complex or  
CC amplification product with presence of hTERT in the sample; and (D)  
CC increasing the proliferation of a vertebrate cell by increasing hTERT  
CC expression; and (E) the use of an agent that causes an increase in cell  
CC vertebrate cell proliferation to create a medicament that inhibits  
CC ageing. A protein preparation of hTERT and the polynucleotide encoding  
CC hTERT can be used in the manufacture of medicaments for inhibiting the  
CC effect of ageing or cancer. Inhibitors of telomerase activity can be  
CC used to treat conditions that are associated with high telomerase  
CC activity. A protein preparation of hTERT can also be used in the new  
CC methods.  
XX Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 19; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLRSRYREVWVPLATFVRRLLGPEGRRLVQDPDKIYETLVAOCLVCMHW 60  
1 MTRAPCPAVRSLRSRYREVWVPLATFVRRLLGPEGRRLVQDPDKIYETLVAOCLVCMHW 60  
61 GSQPPPADLSFHQSLSKELVARVVRQRCERNERNVLAFFGELLNEARGGPPNAAFTSSVR 120  
61 DARPPPAAPSFQVSCVSKELVARVVRQRCERNERNVLAFFGELLNEARGGPPNAAFTSSVR 120  
121 SYLPTNVTIETRVSGAWMLLSRVGDDLLVILLAHCALYLLVPSCAYOYCGSPLYICA 180  
121 SYLPTNVTIETRVSGAWMLLSRVGDDLLVILLAHCALYLLVPSCAYOYCGSPLYICA 180  
181 TTDIWPSVASYRPTPVGRNFTNLRFLLQIKSSSQEQAPKPLALPSGRGTRKHLSTSTS 240  
181 ATQAREPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAPGARRRGSASRS 231  
241 VPSAKARCYVPVREVEGP-----HRQVLTPTSGKSN-VPSPARSEVPTABKOLSSK 292  
232 LPLPKPRRGAAPERTPVGQSWAHFGRTRGSDRGFCVVSAR-----PABEATSLE 286  
293 GKVSIDLSS-GSVCCCHKPSSLSLPPRONAFQLRP-PIETRHFLYSRGDQERLNPSP 350  
287 GALSCTRHSHPSVGRQHAGAPSTSEPRPMDTPCPVYAKTKHFLYSSGD-KEQLAPSP 345  
351 LLSNLQPNITGAARLVEIIFLGSRRPTSGPLCRTHLSRRYQWQMRPLFQQLLVNHAECQY 410  
346 LLSLRPSLTGARRLVETIFLGSRRPMDTPRRLPRLPQRYQWQMRPLFLELLGNHACPY 405  
411 VLLRSHCFRTANQOVTAL-----NTSPHMLDLRLHSSPQVY 452  
406 GVLLKTHCPURAA---VTPAAGVCAKEPQGSVAAPBEEDTPRRLVOLLQKHSPPQVY 462  
453 GFLRACLKVVSASLMTGRNRRFPKLLKFIISLQKYLQSLQELAMWKMVEDCHLRS 512  
463 GFVRACLRILVPPGLWSRHNRRFLNKKFISLGHAKLSQLQELTWKMSVRDCAMLR 522  
513 SPGRKRVPAEHLRLRRIATLFLWLVVYVQLLRSFFYITESTFQKRLFFRKYSVWS 572  
523 SPGVGCVPAEHLRLRRIATLFLWLVVYVQLLRSFFYITESTFQKRLFFRKYSVWS 582  
573 KLOSIGVROHLERVLRELSEVVRHODTLMAPICRLFPKPNGLRPVWNSYSGMT 632  
583 KLOSIGVROHLERVLRELSEVVRHODTLMAPICRLFPKPNGLRPVWNSYSGMT 642  
633 RALGRKQAQHTORLKTFLSMNVYRTKPHLMGSSVLGMNDIYRTVRAFLVRLALDQ 692  
643 RTFRREKRAEHLTSRVKALFVSLNYERARRPGLGASVLGLDDIHRARWTFVLRVRAQDP 702  
693 TPBMVYKADVTGAYDAIPQGLKAVEVYVWNNIRHSESTYCIROYAVVRSDSGOVHKSPFR 752  
703 PPBYLFVYKVDVTGAYDTIPDRUTEVIAIIG-KPNTYCVRIYAVVQAAHGHVTKAFKS 761  
753 QVTLSDLPYMGQFLKHLQSDASALRNSVVBOSISMNESSSLFFDLFLHLSVYK 812  
762 HVSTLTDLPYMGQFVAHLQET--SPLRDAVYIEQSSSLNEASSGLDFVFLRFMCHHAVR 819  
813 IGDECYTQCGIIPQSSSLTLLCSLCFQGMENKLPRAVORDGLLRFVDDLLVTPHLDQ 872  
820 IRGKSYVCCQGIIPQSSSLTLLCSLCFQGMENKLPRAVORDGLLRFVDDLLVTPHLDQ 879  
873 AKTFLSTLVHGVPEYGMINLQKTVNFFVEPOTLGGAAFYQLPAHCLFPWCGLLDITQT 932  
880 AKTFLSTLVHGVPEYGMINLQKTVNFFVEPOTLGGAAFYQLPAHCLFPWCGLLDITQT 939  
933 LEVPCDYSVGAQTSIKTSLTFQSVKAGKTMKNKLSVLRLKCHGLFLDLQVNSLQTVCI 992  
940 LEVQDSYSSYARTSIRASLTFNRFAGKAGNMRRKLFGLVLRKCHSLFLDLQVNSLQTVCI 999

QY 993 NIYKIFLQAYRHACVIOQLPFDQVRKNTLFFLGLIISQASCCYAILKVKNPMTLKAS 1052  
DB 1000 NIYKIFLQAYRHACVIOQLPFDQVRKNTLFFLGLIISQASCCYAILKVKNPMTLKAS 1059  
QY 1053 GS---FPPEAAHMLCYQAFLLKLAHSHVYKCLLGLPRTAQKLLCKLPEATMTILKAAA 1109  
DB 1060 GAAGPLSEAVQWMLCHQAFLLKLTNRVTVVPLGSLRTAQKLLCKLPEATMTILKAAA 1119  
QY 1110 DPALSTDFQITLD 1122  
DB 1120 NPALPSDFKITLD 1132

RESULT 4  
AA43621  
ID AAY43621 standard; Protein; 1132 AA.  
XX AC AAY43621;  
XX DT 26-JAN-2000 (first entry)  
XX DE A human telomerase reverse transcriptase (TRT) polypeptide.  
XX KW Human; telomerase reverse transcriptase; TRT; T lymphocyte activation;  
XX KW dendritic cell; telomerase activity; cancer cell; proliferating cell;  
XX KW immunological destruction; telomerase; cancer; proliferation disease.  
XX OS Homo sapiens.  
XX FN WO950392-A1.  
XX PD 07-OCT-1999.  
XX PF 30-MAR-1999; 93WO-US06898.  
XX PR 31-MAR-1998; 98US-0112006.  
XX PA (GERO-) GERON CORP.  
XX PI Gaeta FCA;  
XX DR WPI: 1999-610845/52.  
XX DR N-PSDB; AAZ30154.  
XX PT Eliciting an in vivo immune response for prevention and treatment of  
XX cancers  
XX PS Claim 3; Fig 1; 26pp; English.  
XX CC The present sequence represents a human telomerase reverse transcriptase  
XX CC (TRT) polypeptide. The protein is used in the method of the invention.  
XX CC The specification describes a method for activating a T lymphocyte,  
XX CC comprising contacting the T lymphocyte with a dendritic cell that  
XX CC expresses a TRT peptide in the context of a MHC class I or MHC class  
XX CC II molecule. The protein causes induction of an in vivo immunological  
XX CC response to telomerase activity. Cancer cells are characterized by  
XX CC expression of endogenous trf gene and the presence of detectable  
XX CC telomerase activity. Therefore, by eliciting a specific immune response  
XX CC to TRT or to TRT-expressing cells, it is possible to selectively target  
XX CC proliferating cells for immunological destruction. The method is used  
XX CC for eliciting an in vivo immune response to telomerase by activating  
XX CC a T lymphocyte, and is useful for prevention and treatment of cancers and  
XX CC other proliferation diseases/conditions.  
XX SQ Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 20; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLRSRYREVWVPLATFVRRLLGPEGRRLVQDPDKIYETLVAOCLVCMHW 60  
1 MTRAPCPAVRSLRSRYREVWVPLATFVRRLLGPEGRRLVQDPDKIYETLVAOCLVCMHW 60

61 GSOPPADLSFTQVSSIKELVARVVOICLERNERNVLAAGFELLNEARGGPPMAFTSSVR 120  
 61 DARPAPAPSPFQVSCLEKELVARVVOICLERNERNVLAAGFELLNEARGGPPMAFTSSVR 120  
 121 SYLNTVITELRVSGAWMLLGRVGDLLVLLAHCALVLPVPSCAVQVCGSPLYOICA 180  
 121 SYLNTVITELRVSGAWMLLGRVGDLLVLLAHCALVLPVPSCAVQVCGSPLYOICA 180  
 181 TTDIMPSVASRPTPVGNRTNLFQOIKSSRQBPAPKPLALPSRGTKRHLSTLSTS 240  
 181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPADGARRRGGASRS 231  
 241 VPSAKARCVPPVPRVEEGP-----HRQVLPSPGSKW-VPSPARSPVETAKDLSKK 292  
 232 LPLPKPRGAPPEPRTVPGGSAWHPGRTGPDGRFCVVSAP-----PAEANSLE 286  
 293 GKVSIDLSS-GSVCKKPSSTSLISPPRQNAQLRP-FIETRLFLXSRGQGRNLNPSF 350  
 287 GALSGRHSHSPSVGRQHAGPSTSRPPRMDTPCPVVAETKHFYSSGD-KQOLRPSF 345  
 351 LLSNLOPNTGARREVEIIFLGSRRPTSGPLCRTHLSRRYQWMPPLFQOLLVNHACQY 410  
 346 LUSLRPSLTGARLVETIFLGSRRPMDTPRRLPRLPORYQWMPPLFLELLGHAQCPY 405  
 411 VRLRSHGRFRTANQOVTDAL-----NTSPHMLDLRLHSPMQVY 452  
 406 GVLLKTHCPRAA---VTPAAGVCAKEFGQSVAAPEBEDTPRLVQLLRQHSPPQVY 462  
 453 GFRLACLCKVVSASLNGTTHRRNRPFFNKKFISLGKYGKLSLOELMWKMYVECHWLS 512  
 463 GFVRLCLRLVPPGLMGSHNRERFLNKKFISLGKAKLSLOELTWKMSVRDCAWLR 522  
 513 SPQKRVPAERHRLERILATFLWMDTVYVOLLRSFFYTESTFOKNRFFVYRKSVM 572  
 523 SPQVGVPAERHRLERILAKFLWLMVYVVELLSFFYVETTFQXNRLFFVYRKSVM 582  
 573 KLSIGVROHLERVLRELSEVQRHODTWLAMPICRLRFPKPNGLRPIVNNYSNGT 632  
 583 KLSIGVROHLERVLRELSEVQRHODTWLAMPICRLRFPKPNGLRPIVNNYSNGT 642  
 633 RALGRKQACHFTQRLKTLFSLNLYERTKPHLGGSVLGMNDIYRTWRAFLVRLALDQ 692  
 643 RTRPREKRAELTSRVALFSLNLYERTKPHLGGSVLGMNDIYRTWRAFLVRLALDQ 702  
 693 TPRMYFVKADVTGAYDAIPQGLVEVANNIRHSESTYCIQYAVVRDSQGVHKSFRR 752  
 703 PPELYFVKVDVTGAYDTIPQDRITEVIAIIPQNTYCVREYAVVQKAAHGVYKAFKS 761  
 753 QVTLSDLPYMGQFLKHLQSDASALRNSVVEIOSISWNESSESLPFFLHFLRHSVVK 812  
 762 HVSTLTDLQPYMRQFVAHLQET--SPLRDVAVIESSSLEASSGLDFVFLRFCHAVR 819  
 813 IGRVCYTCQGGIPQGSSTLCSLCFGDMENKLEAEVORDGLLRLFVDDLLVTPHLQ 872  
 820 IRKSVYVQGGIPQGSSTLCSLCFGDMENKLEAEVORDGLLRLFVDDLLVTPHLQ 879  
 873 AKTFLTLVHGVPEYECMNLQKTVNPPVPGTILGGAAPVQLPAHCLFPMWGLLLDTQT 932  
 880 AKTFLTLVHGVPEYECMNLQKTVNPPVPGTILGGAAPVQLPAHCLFPMWGLLLDTQT 939  
 933 LEVFCDSYGAOTSISLTSFOSVFKAGTKWENKLSVLRLKCHGLFLDLQVNSLQVCI 992  
 940 LEVQSYSSYARTSIRASLTFRGFKAGNMRKLFGLVRLKXLSLFLDLQVNSLQVCI 999  
 993 NIYKIFLLQAYRPHACVQLPFDQRVRKRLATFLGIISSQASCCYAILKVNKPGMTLKAS 1052  
 1000 NIYKIFLLQAYRPHACVQLPFDQRVRKRLATFLGIISSQASCCYAILKVNKPGMTLKAS 1059  
 1053 GS---PPPEAAHWLCVQAFLLKLAHSHVYKCLLGLPRLTAQKLCRKLPEATMTILKAA 1109  
 1060 GAAGPLPSEAVOWLCHQAFLLKLTTHRVTYVPELLGSLRTAQTLSRKLPGTLLTALEANA 1119

QY 1110 DPALSTDEQTILD 1122  
 DB 1120 NPAFSDFKTILD 1132  
 RESULT 5  
 AAY28881  
 ID AAY28881 standard; Protein; 1132 AA.  
 XX AAY28881;  
 AC AAY28881;  
 XX 17-JAN-2000 (first entry)  
 DT Human telomerase reverse transcriptase protein.  
 XX  
 DE Human telomerase reverse transcriptase protein; hTERT; telomerase; hEST2;  
 KW catalytic protein component; cell proliferative capacity; DNA primer;  
 KW telomerase substrate; telomeric DNA synthesis; cell immortality;  
 KW neoplastic phenotype; diagnostic application; prognostic application;  
 KW telomerase related condition; cancer; therapeutic agent;  
 KW telomerase expression; telomerase activity.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Misc-difference 608 /note= "Corresponds to cac codon"  
 FT  
 XX WO9950279-A1.  
 XX 07-OCT-1999.  
 XX 31-MAR-1999; 99WO-US07160.  
 XX 31-MAR-1999; 98US-0052919.  
 XX (GERO-) GERON CORP.  
 PA (UYTE-) UNIV TECHNOLOGY CORP.  
 XX  
 PI Cech TR, Lingner J, Nakamura T, Chapman KB, Morin GB, Harley CB;  
 PI Andrews WH;  
 XX WPI; 1999-610834/52.  
 DR N-PSDB; AA208150.  
 XX  
 PT Antisense polynucleotides for human telomerase reverse transcriptase  
 PT used for diagnosing or treating cancer -  
 XX Claim 2; Fig 2; 31pp; English.  
 CC The present sequence is human telomerase reverse transcriptase protein.  
 CC This is the catalytic protein component of telomerase and is also  
 CC referred to as hEST2. hTERT has the ability to extend a DNA primer that  
 CC functions as a telomeric substrate for telomeric DNA synthesis. This  
 CC correlates with cell proliferative capacity, cell immortality, and the  
 CC development of a neoplastic phenotype. Human TERT antisense  
 CC oligonucleotides are useful for diagnostic or prognostic applications to  
 CC telomerase related conditions, including cancer. They are also useful as  
 CC therapeutic agents, for inhibition of telomerase expression and activity.  
 XX Sequence 1132 AA;  
 SQ  
 Query Match 59.4%; Score 3505; DB 20; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 QY 1 MTRAPRCVAVSLRSRYREVPLATFVRLGPEGRRLVQGPDKIYRTLVACQVCMHW 60  
 DB 1 MPRAPRCVAVSLRSRYREVPLATFVRLGPEGRRLVQGPDKIYRTLVACQVCMHW 60  
 QY 61 GSQPPPADLSFTQVSSIKELVARVVOICLERNERNVLAAGFELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPAPAPSPFQVSCLEKELVARVVOICLERNERNVLAAGFELLNEARGGPPMAFTSSVR 120

121 SYLNTVETLRVSGAMWLLLSRVGDDLLVYLLAHCALVLLVPPSCAYOVCSPLYQICA 180  
121 SYLNTVTDALRGSGAMGLLRVVGDDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA 180  
181 TTDIWPVSASRYRTREVGNTNLRFLQIKSSRQEAQPKPLALPSRGTGKRLHSITSTS 240  
181 ATQARPPHAS-GPRRLG-----CERAWSHVREACVPLGLPAPGARRRGASARS 231  
241 VPSAKARCYVPVRVERGP-----HROVLPTPSGKM-VSPARSPVPTAEXDLISK 292  
232 LPLPKRPRGGAAPERTPVQGSWAHPGRTRGFSDRGFCVWSPAR-----PAEATSLE 286  
293 GKVSPLSLG-GSVCCKHKPSSTLSLSPRQNAQLRP-FLETHFHYLSRGGQERLNSP 350  
287 GALSGRTHSHSVGRQHAGPPSTSPRPDPWDTPCPVVAETHKFLYSSGD-KEQRPSF 345  
351 LLSNLQPNLTGARELVRIIFLGRSPRTSGCLRTHLSRRYQWRELPQCLLVNHAECQY 410  
346 LLSLRPSLTCARLLVETIFLGRSPWPGTPRLPLPQRYWQMRPLFLELLGNHAQCFY 405  
411 VLLSHCRFTANQOVTDAL-----NTSPHLMDLRLHSSPWQVY 452  
406 GVLLKTHCPLRAA---VTPAAGVAREKPGSVAPEEEDTPRRLVQLLRHSSPWQVY 462  
453 GFLRACLKCVSASLWGTNRNERRFFKNLKKFISLKGKLSLQELMWKWKVEDCHWLS 512  
463 GFVRACLRLVPPGLGSRNERNRFLNKKFISLGHAKLSLQELTWKWSVVDCAWLR 522  
513 SPGKDRVPAABHRLRERILATFLFWLMDTVVQLLRSPFFYITESTQKRLFPYKSWVS 572  
523 SPFGVCVPAEAEHRLREELAKFLHMLS VVYVELLSFFYVTTTQKRLFPYKSWVS 582  
573 KLOSIGVQHLEVRVLRSELSOEVRHDDTWLAMPICRLRFIPKPNGLRPIVNMYSMTG 632  
593 KLOSIGICHLKRVQLRELSAEVROHREARPALLTSRLRFIPKPNGLRPIVNMVYGA 642  
633 RALGRKQAQHTORLKTFLFSLMYERTKPHLMGSSVLGMNDIYTWRAFVLRVVALDQ 692  
643 RTFREKRAELTSRVKALFSLVNYERARRPGLLGASVLGLDIDHRAWRTFVLRVAQDP 702  
693 PRMYFVKADVTGADATPOGLKVEVVMNRHSENYCIRQAVVRSDQGVHSPRR 752  
703 PPELVFVVDVTGADVTTPQDLRETVIASLIK-PQNTYCVRRVAVVQKAHGVRAFKS 761  
753 QVTTLSLQPMYGOFLKHLQSDASALNSVYIEQSISMNNESSSLFDFPLHFLHSVYK 812  
762 HVSTLTDLPYMRQFVAHQET--SPLRDAVVEQSLSNEASSGLFDVFLRPMCHAVR 819  
813 IGDRICYTOGIPQGSSTLTLCSLFCGDMENKLPFAVQSDGLLRLFPVDDFLAVTTHLQ 872  
820 IRGKSYVQCGIPQGSILLTLCSLFCGDMENKLPFAGIRDDGLLRLVDDFLVTLPHLTH 879  
873 AKTFSLTLVHGVPEYCGMINLQKTVNPPVPEFGLGGAAPYQLPAHCLFPWCGLLLDTOT 932  
880 AKTFRLTLVRGPEYGVVNLKTVNPPVPEDEALGCTAFVQVPAHGLFPWCGLLLDTRT 939  
933 LEVFCDSYGAQTSITKSLTQSVFKAGTKWRNKLKSLVLRKLCXGLFLDLOVNSLQTVCI 992  
940 LEVQSDYSYARTSIRASLTNRRGFKAGNRNKLFGVLRKLCXSLFLDLOVNSLQTVCT 999  
993 NYKIFLLQAVRFHACVTLQLPDQVRNQLTFFLGIISQAACCYAILKVNKPMGLTKAS 1052  
1000 NYIKILLQAVRFHACVTLQLPDQVQWKNPTFFLRVLSDTASLCSYLKAKNAGMSLGAK 1059  
1053 GS---FPPEAHWLCYQAFLLKLAHSHVYIKLLGPIRLTAQKLCRKLPEATWTILKAA 1109  
1060 GAAGLPSEAVQWLCHQAFLLKTRHRYVTVYVLLGSLRTAQTLQSRKLPGTTLTALEAAA 1119  
1110 DPALSTDFQTLID 1122  
1120 NPALPSDFKTLID 1132

RESULT 6  
AA32090  
ID AA32090 standard; Protein; 1132 AA.  
XX AC AA32090;  
XX DT 17-JAN-2000 (first entry)  
XX DE Human telomerase reverse transcriptase (hTERT).  
XX DE Telomerase reverse transcriptase; human; hTERT; cell proliferation;  
XX KW cancer.  
XX OS Homo sapiens.  
XX FN W09950386-A2.  
XX PD 07-OCT-1999.  
XX PF 31-MAR-1999; 99WO-US07097.  
XX PR 31-MAR-1998; 98US-0052864.  
XX PR 03-AUG-1998; 98US-0128354.  
XX (GERO-) GERON CORP.  
XX PA Morin GB;  
XX PI WPI; 1999-610842/52.  
XX DR N-PSDB; AAZ20279.  
XX DR New catalytic polypeptide and polynucleotide, useful for increasing  
XX PT catalytic activity in a cell -  
XX PS Claim 13; Fig 1; 24pp; English.  
XX CC The present sequence represents human telomerase reverse  
XX CC transcriptase (hTERT). Human telomerase is a target for diagnosing  
XX CC and treating diseases relating to cell proliferation and  
XX CC senescence, such as cancer, or for increasing the proliferative  
XX CC capacity of a cell. A claimed method for increasing the  
XX CC proliferative capacity of a vertebrate cell, especially a human or  
XX CC other mammalian cell, involves introducing into the cell a  
XX CC recombinant hTERT polynucleotide encoding an hTERT variant in which  
XX CC residues 192-323, 200-323, 192-271, 200-271, 222-240, 415-450,  
XX CC 192-323 and 415-450, or 192-271 and 415-450 of the present sequence  
XX CC are deleted. A claimed method of preparing recombinant telomerase  
XX CC involves contacting a recombinant hTERT deletion mutant (as above)  
XX CC with a telomerase RNA component such that the 2 proteins associate  
XX CC to form a complex capable of catalysing the addition of nucleotides  
XX CC to a telomerase substrate. A claimed method for reducing  
XX CC telomerase activity in a cell involves introducing a recombinant  
XX CC polynucleotide encoding an hTERT variant having a deletion of amino  
XX CC acids 192-450, 560-565, 637-660, 638-660, 748-764 or 1055-1071 of  
XX CC the present sequence.  
SQ Sequence 1132 AA;  
Query Match 59.4%; Score 3505; DB 20; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPRCVARSLLRSRYREVWPLATFVRRLQPEGRRLVQPDPKIYRTLVAAQCLVCMHW 60  
DB 1 MPAPRCVARSLLRSRYREVPLATFVRRLQPEGRRLVQPDPAFPAFALVAAQCLVCPW 60  
QY 61 GSQPPPADLSHFQVSSKELVARVQVRLCERNRNVLAFGPELLNEARGPPMATSSVR 120  
DB 61 DARPPPAAPSFQVSCLEKELVARVQLCERGAKNVLAFGFALLDARGGPPFAFTTSVR 120  
QY 121 SYLNTVETLRVSGAMWLLLSRVGDDLLVYLLAHCALVLLVPPSCAYOVCSPLYQICA 180

121	SYLNTVTDALRGSGAWGLLRRVGDVLLVHLLARCALFVLVAPSCAYQVCGPPLYQLGA	180	AY26580 standard; Protein; 1132 AA.
181	TTDIWFSVASYRTPVGRNFTNLRFLOQIKSSROEAPKPLPSRGTKRHLSTSTS	240	AY26580;
181	ATQARPPPHAS-GPERRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS	231	13-SEP-1999 (first entry)
241	VPSAKKACVYPRVVEGPG-----HRQVLPTSGKSW-VPSPARSPVPTAEKDLSSK	292	Human telomerase reverse transcriptase (hTERT) enzyme.
232	LPLKPRPRGAPEPRTFVQGGSWAHFGRTRGSDRGFCVSPAR-----PAEATSLSLE	286	Telomerase reverse transcriptase; TERT; mouse; telomere length assay;
293	GVKSDLSLS-GSVCKKHPSSTLSLSPQNAFQLRP-FIETRHFLYSGDGGERLNPSF	350	immunogen; enzyme; telomerase-mediated DNA replication; human.
287	GALSTRHSHSVGRQHAGPSTSRPPRWDTPCPVYAEKHFLYSSGD-KEQLRPF	345	Homc sapiens.
351	LLSNLOPLTGARRLVEIFLGSRRPTSGCLRTHLSRRYQWRPLFOQLLVNHAQQY	410	WO9927113-A1.
346	LLSLRPSLTGARRLVEIFLGSRRPMDTPRRLPRLPQRYQWRPLFLELLGNHACQY	405	03-JUN-1999.
411	VLLRSHCFRTANQVTDAL-----NTSPHMLDLRLHSSPWQVY	452	25-NOV-1998; 98WO-US25211.
406	GVLTKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSSPWQVY	462	16-MAR-1998; 98US-0042460.
453	GFRLACLKVSASLWGRHNERFFKNLKKFISLGKYGKLSLQELMWKMKVDDCHWLS	512	26-NOV-1997; 97US-0979742.
463	GFVRACLRLVPPGLWGSRHNERFLRTKFIISLGKXKLSLQELTWKMSVRDCAWLR	522	(GERO-) GERON CORP.
513	SPGKDRVPAAEHRLRERILATFLWMDTYVQLRSEFFYITESTFOKNRPLFYKSVMS	572	(YESH) UNIV YESHIVA EINSTEIN COLLEGE.
523	SPGVCVPAAEHRLREELIAFLWMSVYVVELLRSEFFYVTTETFOKNRPLFYKSVMS	592	Allsopp R, Depinho R, Greenberg R, Morin GB;
573	KLOSGVGRHLRERLRELSQBEVRHHQDWTLMPICLRIFPKPGLRPIVNNYSMGT	632	WPI; 1999-347722/29.
583	KLOSGIRCHLKVQLRELSAEVQREAPALLTSLRPIPKDGLRPIVNDYVGA	642	Mouse telomerase reverse transcriptase (mTERT) enzyme proteins and
633	RALGRKQAHQTBQRLKTLFSLMNYRTKPHLMGSSVLGMNDIYRTWRAPVLRALDQ	692	nucleic acids
643	RTFREKRAERLTSVKALFSLVNERARRPGLGASVGLGDDIHRAWRFTVLVRAQDP	702	Disclosure; Fig 3; 135pp; English.
693	TPRMVFKADVTGAYDAPOGKLVEVAVNMIRHSESTVCIROAVVRDSSGQVHKSFR	752	The invention relates to a mouse telomerase reverse transcriptase (mTERT)
703	PBLFVFKVDVTGAYDTPOKRLTEVIAIIT-KPNTYCVRYAVVQKAAHGVKAFKS	761	enzyme. Competitions containing mTERT can be used in telomere length
753	QVTTLSLDQPYMGQFLKHLQSDASALRNSVVIQSISMNESSSLDFLHLRHSVXK	812	assays. Isolated mTERT is useful as an immunogen for the production of
762	HVSTLTDLQPYMRQFVAHLQET--SPLADAVVIRQSSSLNEASSGLFDVFLRFMCHAVR	819	monoclonal or polyclonal antibodies. The method is useful for assessing
813	IGRCYTCQCGIPQSSSLTLLCSLCFGDMENKLPFAEVORDGILLRVDFFLVTPLHQ	872	the degree of purification and identification of new mTERT species, such
820	IRGKSVQCGG:POGSIITLLCSLCYGDENKULFAGIRRDGLLRLLVDFFLVTPLHT	879	as an mTERT allele, homolog or isoform, or to screen for modulators
873	AKTFLLTVHGVPEYGCMLNLOKTVNPPVPGTILGGAAPYQLPAHCLFPWCGLLLDTQT	932	(antagonists and agonists) of telomerase-mediated DNA replication.
880	AKTFLLTVHGVPEYGCMLNLOKTVNPPVPGTILGGAAPYQLPAHCLFPWCGLLLDTQT	939	CC (antagonists and agonists of mTERT can be used to modify the activity of
933	LEVFDYSYATSIKTSITPSQVFAKATWKNLLSVLRUKCHGLFLDLOVNSLQTVCI	992	CC other telomerase enzymes such as human TERT (hTERT). The present sequence
940	LEVQSDYSYATSIKTSITPSQVFAKATWKNLLSVLRUKCHGLFLDLOVNSLQTVCT	999	CC represents a human TERT enzyme.
993	NYIKFLQAVRFHACVQLPQDQVRKNTLFFLIGIISQASCCVAILKVNKPMWTLKAS	1052	Sequence 1132 AA;
1000	NYIKILLQAVRFHACVQLPQDQVRKNTLFFLIGIISQASCCVAILKVNKPMWTLKAS	1059	Query Match 59.4%; Score 3505; DB 20; Length 1132;
1053	GS---FPPPAHWCYQAFLLKLAHSHVYIKCLLGLPLTAQKL:CRKLPEATMTLKAAA	1109	Best Local Similarity 62.4%; Pred. No. 0;
1060	GAAGPLPSEAVQWCHQAFLLKLTSHRVTVYVPLLSLTAQQLSRKLPGLTTLTALEAAA	1119	Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;
1110	DPALSTDFOTILD 1122		1 MTRAPCPAVRSILRSRYREVWPLATFVRRLGPEGRRLVQPGDPKTYRITLVAQCLVCMHW 60
1120	NPALPSDFKTIID 1132		1 MPAPRCVAVRSILRSRYREVWPLATFVRRLGPEGRRLVQPGDPKTYRITLVAQCLVCMHW 60
			61 GSQPPADLSHQVSSIKELVARVQVQLCERNERNVLAFCFELLNEARGGPPMAFTSSVR 120
			61 DARPPPAAPSFRQVSCIKELVARVQVQLCERNERNVLAFCFELLNEARGGPPMAFTSSVR 120
			121 SYLNTVTETLRVSGAMLLLRVGGDDLLVLLAHACALYLLVPPSCAYQVCGSPLYQICA 180
			121 SYLNTVTETLRVSGAMLLLRVGGDDLLVLLAHACALYLLVPPSCAYQVCGSPLYQICA 180
			181 TTDIWPVSASRYRTPVGRNFTNLRFLOQIKSSROEAPKPLPSRGTKRHLSTSTS 240
			181 ATQARPPPHAS-GPERRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS 231
			241 VPSAKKACVYPRVVEGPG-----HRQVLPTSGKSW-VPSPARSPVPTAEKDLSSK 292
			232 LPLKPRPRGAPEPRTFVQGGSWAHFGRTRGSDRGFCVSPAR-----PAEATSLSLE 286
			293 GKVSDLSLS-GSVCKKHPSSTLSLSPQNAFQLRP-FIETRHFLYSGDGGERLNPSF 350

```

287 GALSCTRHSHPVSGRQHAGPPTSRPRPWRDTCPPVVAETKHFYSSGD-KEQLRPSF 345
351 LLSNLQPNLTGARRLVEIIFLGSRPTSGCLTRHLSRRYQMRPLFQQLVNHAEQY 410
346 LLSLRPSLTGARRLVEIIFLGSRPMPGTPRRLPRLPQRYQMRPLFELLGNHAQCFY 405
411 VRLLSHCRFTANQOVDAL-----NTSPHLMDLRLHSSPWQY 452
406 GYLLKTHCPLRAA---VTPAGVCAREKPOGSAVAPEEDTDPRLVQLLRHSSPWQY 462
453 GFLRACLKVVYSASLWGRHNERFFKNLKKFISLGKYKLSLQELMWMKVEDCHWLR 512
463 CFVACLRLVPPGLANGSRNRRFLENTKFIISLGHAKLSLQELTWMSVEDCAWLR 522
513 SPGRDVPAAERLRERILATFLFWLMDTVVOLLRSFFYITESTQKRLFPYKSVWS 572
523 SPGVGCPAAERHREELAKFLHLMVSVVVELLSFFYVTTTQKRLFPYKSVWS 582
573 KLOSTGVQHLERVRRLRELSOEVRHHDWTWAMPICRLRFIPKPNGLRPIVNMYSMT 632
583 KLOSTIGIRHUKRVQLRELSAEVRQHRARPALLTSRLRFIPKPNGLRPIVNMVYGA 642
633 RALGRKQAOHFTQRLKTLFSLMNYERTKPHLMGSSVLGMNDIYRTWAFVLRVADQ 692
643 RTFREKRAELTSRKVALFSVLYNERARRPGLLGASVLGLDDIHEAWRTFVLRVAQDP 702
693 TPRLYFKADVTGAYDAIPQGLVEVVMIRHSESTYCIQYAVVRDSQGVHKSFR 752
703 PPELFYKVDVTGAYDITPQRLTEVIASIIK-PQNTYCVRRYAVVQKAHGVKAFKS 761
753 QVTTLSLQPMYGOFLKHLQSDASALRNSVYIQSISMNESSSLFDFPLHFLRHSVYK 812
762 HYSTLTDLQPMYQFVAHQET--SPLEDAVVIQSSSLNEASSGLFDVFLRPMCHAVR 819
813 IDRCYTCQGIPOGSSUSTLLCSLFGDMENKLPABVQRDGLLRFVDDFLVTHLQ 872
820 IAGKSYVQCQSIPOGSSILLTLLCSYCDMENKLPAGIRRDGLLRLVDDFLVTHL 879
873 AKTFTSLTVHGVPEYGCMMINLQKTVNPPVPGTGGAAVQLPAHCLFPWCGLLDTOT 932
880 AKTFTLTVRGVPEYGCVMNLRKTVNFPVEDEALGGTAFVQMPAHGLFPWCGLLDT 939
933 LEVPCDYSYGAQTSIKTSLTQSVFKAGTKWRNKLVLRLKHGLFLDLQVNSLQTVCI 992
940 LEVQSDYSYARTSIRASILTNRGFKAGNRMRKLFGLRLKCHSLFLDLQVNSLQTVCT 999
993 NIYKIFLLOAYRFHACVQLPDPQVRKNTLFFELGIISSOASCCYAILKVNPMGLTKAS 1052
1000 NIYKILLQAYRFHACVQLPDPHQVKNPTFFLRVSDTASLCYSLKAKNAGMSLGAK 1059
1053 GS---FPPPAHWCYQAPLLKLAHSHVYIKLLGPLRTAQKLCRLKPLATMTILKAAA 1109
1060 GAAGPLPSEAVQWLCHQAFLLKLTSHRVTVYVPLGLSLRTAQQLSRKLPGLTTLTAA 1119
1110 DPALSTDFQTIID 1122
1120 NPALPSDFXTIID 1132

```

SULT 8  
W90251  
AAW90251 standard; Protein; 1132 AA.  
AAW90251;

24-MAY-1999 (first entry)

Human catalytic telomerase sub-unit protein.

Human; catalytic telomerase subunit; therapy; diagnosis; htc; assay; modulator; treatment; inhibit; cellular disorder; death; defect; cancer; ageing; antisense; neoplastic cell; telomerase-related condition; tumour cell.

```

XX OS Homo sapiens.
XX PN WO9859040-A2.
XX PD 30-DEC-1998.
XX PF 09-JUN-1998; 98WO-EP03468.
XX PR 14-APR-1998; 98DE-1016496.
XX PR 20-JUN-1997; 97DE-1026329.
XX PR 26-MAR-1998; 98DE-1013274.
XX PA (FARB ) BAYER AG.
XX Hagen G, Siegmund H, Weichel W, Wick M, Zubov D;
XX WI; 1999-081276/07.
XX DR N-PSDB; AAV72117.
XX PT New catalytically active subunit of human telomerase - used in the
XX PT modulation of telomerase activity, particularly for treating cancer
XX PS and ageing
XX PS Claim 2; Fig 2; 76pp; German.
XX CC This sequence represents a novel human catalytic telomerase sub-unit
XX CC (htc). This protein can be used in screening assays to identify
XX CC modulators of telomerase and to treat or inhibit cellular disorders,
XX CC death, defects and/or other pathological processes involving telomerase,
XX CC particularly cancer and ageing (also suitable for this are agents that
XX CC stimulate, inhibit or mimic the activity of the subunit). Antisense
XX CC nucleic acids inhibit telomerase action (by binding to specific mRNA),
XX CC particularly in neoplastic cells and may be expressed in vivo. Antibodies
XX CC and fragments of the protein, used as probes or primers, are used to
XX CC diagnose telomerase-related conditions (especially neoplasia) by (i)
XX CC detecting abnormal levels of the subunit protein in body fluids or
XX CC tissues or (ii) by measuring the amount of the encoding nucleic acid.
XX CC Expression of the nucleic acid encoding the subunit mRNA is confined to
XX CC tumour cells, in contrast to the ubiquitous expression of the telomerase
XX CC RNA subunit.
XX SQ Sequence 1132 AA;
Query Match 59.4%; Score 3505; DB 20; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;
QY 1 MTRAPCPAVRSLLRGRYREVWPLATFVRLGPEGRRLVQPGDKIYRTLVAQCLVCMHW 60
DB 1 MPAPRCRAVRSLLRSHYREVPLATFVRLGPGQRLVQRGDPAPAFRALVAQCLVCPW 60
QY 61 GSQPPPADLSHFQVSSILKELVARVQRLCERNERNVLAFGPELLNEARGGPPMATSSVR 120
DB 61 DARPPPAAPSFQVSCVCLKELVARVQLCERGAKNVLAFGFALLDGCARGGPEAFTTSVR 120
QY 121 SYLNTVIEILRVSGAMWLLSRVGDLLVLLAHALYLLVPBSCAYOVCGSPYQICA 180
DB 121 SYLNTVTDALRGSGAMGLLRVGDVLLHLLARCALFVLVAPSCAYOVCGPPYQLGA 180
QY 181 TTDIWPVSASVYRPTREVGENTFLRFLQIKSSSSRQEAQKPLALPSRGTKRHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRELG-----CERAWNHSVREAGVPLGLPAGARRRGGASRS 231
QY 241 VPSAKARCPVPVRVEEGP-----HRQVLPFPGSKSW-VPSPARSPVPTAEKDLSSK 292
DB 232 LPLPKRPRGAPEPFTFVGQGSWAHPGTRGPDGFCVVSAPR-----PAEATSLSE 286
QY 293 GKVSIDLSS-GSVCKKPSSTGILLSPPRONAFQLRP-FIETRHFLYSRGDGQERLNPSF 350
DB 287 GALSCTRHSHPVSGRQHAGPPTSRPRPWRDTCPPVVAETKHFYSSGD-KEQLRPSF 345
QY 351 LLSNLQPNLTGARRLVEIIFLGSRPTSGCLTRHLSRRYQMRPLFQQLVNHAEQY 410

```





15-FEB-2001; 2001WO-US05143.

15-FEB-2000; 2000US-0182685.

15-FEB-2001; 2001US-0182685.

(REGC ) UNIV CALIFORNIA.

Zanetti M;

WPI; 2001-536552/59.

Vaccine for initiating and enhancing a cytotoxic T lymphocyte response, for treating cancers or tumors or for inducing immune response against tumors, comprises a telomerase reverse transcriptase peptide

Disclosure; Fig 5; 52pp; English.

The present sequence is that of human telomerase reverse transcriptase (hTERT). The sequence was analysed for 9-mer peptide sequences containing known binding motifs for the human leukocyte antigen HLA-A2.1 molecule. From an initial panel of about 30 candidate peptides, 2 sequences, denoted p540 (see AAB82772) and p865 (see AAB82773), were examined. The majority of healthy individuals as well as patients with prostate cancer immunised in vitro against these 2 HLA-A2.1 restricted peptides developed hTERT-specific cytotoxic T lymphocytes (CTL). The cancer patients' CTL specifically lysed a variety of HLA-A2+ cancer cell lines such as prostate, breast, colon, lung and melanoma, demonstrating immunological recognition of endogenously-processed hTERT peptides. In vivo immunisation of HLA-A2.1 transgenic mice generated a specific CTL response against both hTERT peptides. The induction of CTL responses in vitro and in vivo, and the susceptibility to lysis of tumour cells of various origins by hTERT CTL suggest that hTERT could serve as a universal cancer vaccine for humans. Thus, a claimed universal vaccine for treating tumours of any origin comprises at least 1 hTERT peptide in an amount effective for initiating and enhancing a CTL response against cancer cells. The peptide is 7-15 amino acid residues in length and may be modified to enhance binding to the major histocompatibility complex. Also claimed is a method for inducing and enhancing a CTL response against cancer cells, involving harvesting blood leukocytes, pulsing with hTERT, and contacting cancer cells with the pulsed leukocytes. A method for targeting CTL to tumour cells is also claimed, and involves administering a hTERT peptide to a mammal, especially a cancer patient.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
Best local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
Y 1 MTRAPRCVAVRSLLRSRYREVWPLATVRLGPEGRLVQGDPKIYRTLVAAQCLVCMHW 60  
b 1 MTRAPRCVAVRSLLRSRYREVWPLATVRLGPEGRLVQGDPAFRAVLAQCLVCPW 60  
Y 61 GSQPPADLSHQVSSLSKELVARVQBLCEBENRVLAQFELNARGPPMAFTSSVR 120  
b 61 DARPPPAAPSPFQVSKELVARVQBLCEBENRVLAQFELNARGPPMAFTSSVR 120  
Y 121 SYLPTNTVETLTVSGAMWLLSRVGGDLLVYLLAHACALVLLVPPSCAYQVCGSPLYQICA 180  
b 121 SYLPTNTVETLTVSGAMWLLSRVGGDLLVYLLAHACALVLLVPPSCAYQVCGSPLYQICA 180  
Y 181 TTDIWPVSASVYRTPVGRNFTNRLFQIKSSROEAPKPLALPSRGTRHLSSTTS 240  
b 181 ATQARPPPHAS-GPFRRLG-----CERAWNSHREAGVPLGLFAPGARRGSSASRS 231  
Y 241 VPSAKACVYRVEEGP-----HQVLPDPSGKSW-VPSPARSPEVPTAEKDLSSK 292  
b 232 LPLPKRPRGAAPEPFTVQGGSHAPGTRGDSRGCVCVSPAR-----PAEATSL 286  
Y 293 GKVSLSLS-GSVCKKHKPSSTLSLSPRONAFQLRP-FIETRHFLYSGDGOERLNPSF 350

Db 287 GALSTRSHSVSGRQHAGPSTSRPRPMDTFCPPVYAEATKHFLYSSGD-KEQLRPSF 345  
QY 351 LLSNLQPNLTGARRLVEIFLGSRPRTSGPLCRTHRLSRRYQMRPLFOQLLVNHAECQY 410  
Db 346 LLSLSRPSLTGARRLVEIFLGSRPRTSGPLCRTHRLSRRYQMRPLFOQLLVNHAECQY 405  
QY 411 VLLRSHCHCFRTANQQVTDAL-----NTSPHLMDLRLRLHSSPMQVY 452  
Db 406 GVLLKTHCPRLAA---VTPAAGVCAREKPKQSVAAPEEDTDPRLLVQLLRQSSPMQVY 462  
QY 453 GFRLACLKVVYASLWGTNRHNERFFKNLKKFISLGYKLSLQELMMKMKVEDCHWLRS 512  
Db 463 GFVRACLRELVPGLGSRHNERFLNKKFISLGYKLSLQELMMKMKVEDCHWLRS 522  
QY 513 SPGRVPAAEHRLERILATFLWMDTVVQLRFFYFTESTFQKNLFFYRKSVM 572  
Db 523 SPGVGCPAAEHLREELAKFLWMSVYVVELRFFYFTESTFQKNLFFYRKSVM 582  
QY 573 KLSIGVROHLERVELRSLQSEVRRHODTWLAMPICRLRPIPKPGLRPIVNMYSMGT 632  
Db 583 KLSIGVROHLERVELRSLQSEVRRHODTWLAMPICRLRPIPKPGLRPIVNMYSMGT 642  
QY 633 RALGRKQKHQFTQBLKILFSLMNLVYRTKPHLMGSSVGLGNDIYRTWRAFLVRAALDQ 692  
Db 643 RTFRREKRAELTSRVKALFSLVNYERARRPGLLGASVGLDDIHRARWTFVLVRADQP 702  
QY 693 TRPMYFVKADVTGAYDAIPQGLVEVAVNMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
Db 703 PELYFVKVDTGAYDTTPODELTEVIAIILK-PQNIYCYRYAVVQKAAHGHVYKAFKS 761  
QY 753 QVTTLSDLQPYNGQFLKHLQSDASALRNSVVIQSSISMNESSSLDFPLHFLRHSVVK 812  
Db 762 HYSTLTDLPYMRQFVAHLQET--SPLRDAVVVQSSSLNEASSGLDFVFLRFMCHAVR 819  
QY 813 IGRDCTVQCGIPQGSSTLILCSLFCGMENKLPFAEVRDGLLRFVDDFLVTPHLQ 872  
Db 820 IRKSVVQCGIPQGSSTLILCSLFCGMENKLPFAEVRDGLLRFVDDFLVTPHLQ 879  
QY 873 AKTFLSTLVHGVPEYGCMMINLQKTVYVFPFPGTLGGAAPYQLPAHCLFPWCGLLDTQT 932  
Db 880 AKTFLSTLVHGVPEYGCMMINLQKTVYVFPFPGTLGGAAPYQLPAHCLFPWCGLLDTQT 939  
QY 933 LEVFCYSYAGTSIKTSITFOSVFKAGTWENKLLSVLRKCHGLDLOVNSLQTVCI 992  
Db 940 LEVFCYSYAGTSIKTSITFOSVFKAGTWENKLLSVLRKCHGLDLOVNSLQTVCI 999  
QY 993 NIYKIFELQAYRFHACVIOLEPQDQVRVKNLTPFLGIISSQASCCYAILKVNPGMTLKA 1052  
Db 1000 NIYKIFELQAYRFHACVIOLEPQDQVRVKNLTPFLGIISSQASCCYAILKVNPGMTLKA 1059  
QY 1053 GS---PPEAAHNLVQAFLLKLAHSHVYIKCLQPLRTAQKLLCRKLPEATMTILKAAA 1109  
Db 1060 GAAGPLFSEAVQWLCHQAFLLKLRHRTVYVPLLSGLRTAQKLLCRKLPEATMTILKAAA 1119  
QY 1110 DPALSTDFOTILD 1122  
Db 1120 NPALPSDFKTILD 1132

RESULT 10

AAB99930

ID AAB99930 standard; Protein; 1132 AA.

XX AAB99930;

DT 26-SEP-2001 (first entry)

XX Human telomerase protein sequence SEQ ID NO:31.

XX Differentiation; heart muscle cell; cytokine; transcription factor;  
KW Proliferation; surface antigen; heart disease; cardiomyocyte;  
KW Bone marrow; umbilical blood cell; heart muscle degeneration;

myocardial infarction.

Homo sapiens.

WO200148150-A1.

05-JUL-2001.

02-NOV-2000; 2000WO-JP07741.

28-DEC-1999; 99JP-0372826.

28-FEB-2000; 2000WO-JP01148.

(KYOW ) KYOWA HAKKO KOGYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;

Yamada Y;

WPI; 2001-425655/45.

N-PSDB; AAH44366.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease

Claim 146; Page 137-141; 187pp; Japanese.

The present invention describes cells originating in bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. Also described are: (1) cardiomyocytes produced by the differentiation of the cells; (2) a method for carrying out the differentiation into cardiomyocytes, regulated by a promotional and/or inhibitory factor; (3) a method for the differentiation of the cells into cell types other than cardiomyocytes; (4) drug compositions promoting the formation of heart muscle and regeneration of heart tissue which contain the cells; (5) a method for the production of antibodies which recognise the cells, especially antibodies which recognise a surface antigen on the cells; (6) a method for screening factors which promote the proliferation of the cells; (7) a method for immortalising the cells by expressing telomerase in them; (8) drug compositions for the treatment of heart disease which contain the immortalised cells; and (9) cell-free supernatant from the culture of the cells and its use in promoting their differentiation into cardiomyocytes. The cells are used in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction and in the study of cardiomyocyte differentiation. AAH44351 to AAH44409 and AAB99915 to AAB99935 represent sequences used in the exemplification of the present invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MPRAPRCVAVSLRSRYREYVWPLATFVRRLGPRRLVOPGPKIVRTLVAQCLVCMHW 60

1 MPRAPRCVAVSLRSRYREYVWPLATFVRRLGPRRLVOPGPKIVRTLVAQCLVCMHW 60

61 GSQPPADLSHQVSSKELVARVQRLCERNERNVLAQFELLNEARGPPNMAFTSSVR 120

61 DARPPPAAPSFQVSCLELVARVQRLCERNERNVLAQFELLNEARGPPNMAFTSSVR 120

121 SVLPNTVTLTRVSCAMWLLSRVGGDLLVLLAHCALYLLVPPSCAYQVCGSPFLVQICA 180

121 SVLPNTVTLTRVSCAMWLLSRVGGDLLVLLAHCALYLLVPPSCAYQVCGSPFLVQICA 180

181 TTD1WPSVASYRTPRVPGRNFTNLRFLQIKSSSRQEPAPKPLALSRGKTRHLSLTSTS 240

181 ATOARPPPHAS- GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGSSASRS 231

241 VPSAKKARCYVPVRVEEGP-----HRQVLPTPSGKSW-VPSPARSPFVPTAKDLSSK 292

232 LFLPKPRRGAAPERTFVQGSWAHFGRTRGFSDRGFCVSPAR-----PAAEATSLE 286

QY 293 GKVSDDLGLS-GSVCKKPKSSSTLSLSPRONAFQURP-FIETRHFLYSKGDQGERINPSF 350  
 DB 287 GALSGRSHSPSVGRCHAGAPPSTSRPPRMDTPCPVYAEKHFELYSKGD-KEQLRPSF 345  
 QY 351 LLSNLQNLGARRLVEIIFLGSRPRTSGPLCRTHLSRRYQWMPPLFOQLLVNHAEOY 410  
 DB 346 LLSLRPSLTGARRLVETIFLGSRPMPGTPRLPRLPORYQWMPPLFLELLGNHAQCY 405  
 QY 411 VRLRSHCRFRATANQVTDAL-----NTSPPHLMDLLRLHSSPWQY 452  
 DB 406 GVLLKTHCPRLAA---VTPAAGVCAREKPKQGSVAAPEDTDPRRLVQLLRQHSSPWQY 462  
 QY 453 GFPRACLCVKVVSASLWGTNRRPPFKNLKPTISLGYKGLSLQELMMKMKVEDCHWLS 512  
 DB 463 GFVACURRLVPPCLWGSRRNRRFLNKKTKISLGHAKLSLQELTWKMSVRDCAWLR 522  
 QY 513 SPGRDVRPAAAEHLRERILATFLWLMDTYVQLLRSFFYITESTFQKRLFFYKSVMS 572  
 DB 523 SPGVGCVPAAEHLRERILAKFLHMLMSVVVVELLSFPYVTTETTFQKRLFFYKSVMS 582  
 QY 573 KLSIGVROHLERVRRELSQEVRRHQDTWLMPICLRLRFPKNGLRPIVNMYSMGT 632  
 DB 583 KLSIGVROHLKRVQLRELSEAEVRQREARPALTSRLRFIPKPDGLRPIVNMYYVGA 642  
 QY 633 RALGRKQAQHFQRLKTLFSLMLNYERTKHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
 DB 643 RTFRERKGAERLTSRVKALFSLVNERARRPGLGASVLGLDDIHRWRTFLVRAQDP 702  
 QY 693 TPRMYFKADVGTGAYDAIPGKLVVVVAMRHSSSTYCIROYAVVRDRSQGVHKSFR 752  
 DB 703 PPELYFYKVDVTGAYDTIPQDRLTEVIASIIK-PQNTYCVREYAVVQKAAHGVKAFKS 761  
 QY 753 QVTLSLQPYMGQFLKHLQSDASALRNSVIEQSI-SMNESSSLDFFLHFLRHSVVK 812  
 DB 762 HVSLTDLQPYMQFVAHQET--SPERDAVIEQSSSLNEASSGLFVFLRPMCHAVR 819  
 QY 813 IGDRYTCQCGIPQSSSLSTLCSLCFQGMENKLFABVQVORDGLLLRFVDDFLLVTPHLQ 872  
 DB 820 IRGKSYVQCGIPQSSSLSTLCSLCYQGMENKLFAGIRRDGLLLRLVDDFLLVTPHLTH 879  
 QY 873 AKTSLSLVHGVPEYCMINLQTVNFPVEPCTGGAPYOLPAHCLFPWCGLLDITOT 932  
 DB 880 AKTFLRLTVRGVPEYGCVMNLRKTVNFPVEDEAGTAFVQMPAHGLFPWCGLLDITRT 939  
 QY 933 LEVPCDYSVQAQTSIKTSTTFQSVFKAGTMRNKLISVLRLKCHGLFLDLQVNSLQTVCI 992  
 DB 940 LEVQSDYSVYARTSIRASLTFRNRFKAGRMNRKLFGLVRLKCHSLFLDLQVNSLQTVCT 999  
 QY 993 NIYKIFLQAYRFHACVILQFPDQVRVKNLTFFLGLISSQASCCVAILKVPKPGMTLKAS 1052  
 DB 1000 NIYKIFLQAYRFHACVILQFPDQVRVKNLTFFLGLISSQASCCVAILKVPKPGMTLKAS 1059  
 QY 1053 GS---PPEAAHMLCYQAFLLKLAASHVYKCLLGLPLRTAQKLLCKRKLPEATMTILKAA 1109  
 DB 1060 GAQPLSEAVOMLCHQAFLLKLTTRVTVYVPLGSLRTAQQLSRKLPGLTILTALEAA 1119  
 QY 1110 DPALSTDFQILD 1122  
 DB 1120 NPALPSDFKILD 1132

RESULT 11

AAG64329

ID AAG64329 standard; Protein; 1132 AA.

XX AAG64329;

AC AAG64329;

DT 24-SEP-2001 (first entry)

XX Human protein #2.

DE Angiogenesis; cardiant; cell differentiating agent; bone marrow;

KW

heart muscle cell; heart disease; human.

Homo sapiens.

W0200148149-A1.

05-JUL-2001.

28-FEB-2000; 2000WO-JP01148.

28-DEC-1999; 99JP-0372826.

(KYOW ) KYOWA HAKKO KOGYO KK.

Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K;

WPI; 2001-418252/44.

N-PSDB; AAH49601.

New adult bone marrow-originated cells capable of differentiating into heart muscle cells, applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration .

Disclosure; Pages 128-134; 158pp; Japanese.

The present invention relates to cells isolated from bone marrow, which are capable of at least differentiating into heart muscle cells. The cells are applicable as remedies for various heart diseases particularly with damaged heart muscle accompanying degeneration. The present sequence was used to illustrate the present invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 1 MTRAPCPAVRSLLRGRYREVWPLATFVRLGPEGRRLVQPGDKIYTLVAQCLVCMHW 60  
 1 MPRAPCRVRSLLRSHYREVPLATFVRLGPGQWRLVQGRDPAAFRALVAQCLVCPW 60  
 61 GSOPPPADLSFHVSSIKELVARVQRLCERNVLAFLGPELLNEARGPPMAFTSSVR 120  
 61 DARPPPAAPSFRQVCLKELVARVQRLCERAKVLAFLGALLDARGGPEAFTSSVR 120  
 121 SYLFTVETIRVSGAMLLSRVGDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
 121 SYLFTVTDALRGSGAWGULLRRVGDVLLVHLLARCALPVLVAPSCAYQVCGPPLYQLGA 180  
 181 TTDIWPSVSASYPTRPVGNFTNRLFLQIKSSSRQEPKPLALPSSGTRKHLSTTS 240  
 181 ATQARPPPHAS-GRRRLG-----CERAMNHSVREAGVFLGPAAPARRRGSASRS 231  
 241 VPSAKKARCPVPRVEBGP-----HROVLTPPSGKSW-VPSPARSPEVPTAEKDLSSK 292  
 232 LPLPKRPRRGAPEPEPTVPGQSWAHPCGRTRGSPDRGFCVVSAPR----PAEATSLE 286  
 293 GKVEDLSLS-GSVCCKKPSSTLLSPPQNAFQLRP-FIETRHFLYRGDQGERLNSPF 350  
 287 GALSGRHSHPSVGRQHAGPSTSRPPRPMWDTPCPPVVAETKHFLYSSGD-KEQLRPSF 345  
 351 LLSNLQNLGTARRLVEIIFLGRSPRTSGPLCTHLSRRYQWMPRLFOQLLVNHAECQY 410  
 346 LLSLRSLGTARRLVEIIFLGRSPWPGTFRPLPRLPQRYQWMPRLFLELLGNHAQCPY 405  
 411 VLLRSHCRFTANQOVTDAL-----NTSPPHLMDLLRLHSSPWQVY 452  
 406 GVLLKTHCPLEAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSSPWQVY 462  
 453 GFLRACLCVKVSAISLWTRNRRFFKLLKFKSLGKYKLSLOELMWKVEDCHWLRS 512  
 463 GFVRACLRRLVPPGLWGRSHRNERFLNTKKFISLGHAKLSLQELTWKMSVRDCAWLRR 522

QY 513 SPGRDVPAAEHLRLERILATFLFWLMDTVVQLLSFFVITESTFQKRLPFYKSVMS 572  
 DB 523 SPGVCVPAAEHLRLEREILAKFLHMLSVTVVLLRSFFVITETTFQKRLPFYKSVMS 582  
 QY 573 KLSIGVRLHLRLERLRLSRELSQEVRRHQDTWLAMPICRLAFIPKPNGLRPIVNMYSYMG 632  
 DB 583 KLSIGIRQLHLKRVQLRELSEAEVRQREARPAALLTSRLAFIPKPDGLRPIVNDYVGA 642  
 QY 633 RALGRKQACHFTORLKTFLSMLNRYERTKPHLMSGSVLGMNDIYRTWAFVLRVRLDQ 692  
 DB 643 RTFRREKKAELTGRVAFVLMYERARRPGLGLASVLGLDDIHRARTFVLRVRAQDP 702  
 QY 693 TPRMYFVADYTGADYDAIPQGLVEVVANMRHSESTYCIQYAVVRRDSQGVHKSFR 752  
 DB 703 PPELYFVVDVTDGADYDIPQRLTEVIASIIK-PONTYCVRRYAVVQKAAGHVRKAFKS 761  
 QY 753 QVTLSLDQPMQGLKHLQDSASALRNSVIEQSISSNMESSSSLFDFELHRLHRSVVK 812  
 DB 762 HVSTLTDLQPMQFVAHLQET--SPLRDVAVVIEQSSSLNEASSGLFDFELRMCHHAVR 819  
 QY 813 IGDRCYTCQCGIPQSSSLSTLLCSLCFQDMENKLFABVQDGLLLRPVDDFLLVTPHLQ 872  
 DB 820 TRGKSYVQCGIPQSSSLSTLLCSLCYQDMENKLFAGIRDRGLLLRLVDDFLLVTPHL 879  
 QY 873 AKTFLSTLVHGVPGYGMINLQKTVVNFPEPPTLGGAAFYQLPAHCLFPWCGLLDTQT 932  
 DB 880 AKTFLRTLVRGPGYGVGNLRTKTVVNFPEDEALGTAFAVQMPAHGLFPWCGLLDTT 939  
 QY 933 LEVFCDSYGAQTSIKTSLTFSQSVKAGKTMNKLLSVLRKLGHLFLDLQVNSLOTVC 992  
 DB 940 LEVQSDYSYARTSIRASLTNRGFKAGRNWRKLFVLRKLGHLFLDLQVNSLOTVCT 999  
 QY 993 NIYKIFLLQAYRFHACVQLPFDQVRKNTLTFGIITSSOASCYAILKVNFGMTLKAS 1052  
 DB 1000 NIYKILLQAYRFHACVQLPFDQVQWKNPTFFLRVSDTASLSCYSLKAKNAGMSLGAK 1059  
 QY 1053 GS---FPEAAHNLQYAFILKLAHSHVYIKLIGLPLRTAQKLCRKLPEATWTILKAAA 1109  
 DB 1060 GAAGFLPSEAVQWICHQAFLLKLTTHRVTVYVPLLSURTAQTQLSRKLPGTTTLTALFAAA 1119  
 QY 1110 DPALSTDFQTILD 1122  
 DB 1120 NPALPSDFKTILD 1132  
 RESULT 12  
 AAG64859  
 ID AAG64859 standard; Protein; 1132 AA.  
 XX AAG64859;  
 XX AC  
 XX DT 21-SEP-2001 (first entry)  
 XX DE Heart muscle cell differentiation related protein SEQ ID NO: 31.  
 XX KW Heart muscle cell; human; cell differentiation; heart disease.  
 XX OS Homo sapiens.  
 XX PN W0200148151-A1.  
 XX PD 05-JUL-2001.  
 XX PF 27-DEC-2000; 2000WO-JP09323.  
 XX PR 28-DEC-1999; 99JP-0372826.  
 XX PR 28-FEB-2000; 2000WO-JP01148.  
 XX PR 02-NOV-2000; 2000WO-JP07741.  
 XX PA (KYOW ) KYOWA HAKKO KOGYO KK.  
 XX PI Umezawa A, Hata J, Fukuda K, Ogawa S, Sakurada K, Gojo S;  
 PI Yamada Y;

WPI; 2001-425656/45.  
N-PSDB; AAH468235.

Cells capable of differentiating into cardiomyocytes and originating in bone marrow or umbilical blood cells for study of cardiomyocyte differentiation and treatment of heart disease -

Claim 87; Page 143-147; 183pp; Japanese.

The present invention provides cells originating in the human bone marrow or umbilical blood cells which are capable of differentiating into cardiomyocytes. These cells are useful in the treatment of diseases involving heart muscle degeneration, such as myocardial infarction, and the study of cardiomyocyte differentiation. The present sequence is a protein described in the exemplification of the invention.

Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 22; Length 1132;  
est Local Similarity 62.4%; Pred. No. 0;  
atches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVSLRSYREYVWPLATVRLGEGRRLLVQDPDKIYRVLVACCLVCMHW 60  
1 MPRAPCRVAVSLRSYREYVWPLATVRLGEGRRLLVQDPDKIYRVLVACCLVCMHW 60  
61 GSOPPPADLSPHYSSKLKELVARVVOQLCERNERNVLAFGELLNEARGPPMAFTSVR 120  
61 DARPPPAAPSPQVSCHELVARVVOQLCERNERNVLAFGELLNEARGPPMAFTSVR 120  
121 SYLNTVETLRVSGAMWLLLSRVGDDLLVYLALHACALYLLVPSCAYQVCGSPYQICA 180  
121 SYLNTVTDALRGSGAMWLLLSRVGDDLLVYLALHACALYLLVPSCAYQVCGSPYQICA 180  
191 TTDWPSVASRYTRPVGNFTNLRFLQIKSSRQEARPKPLALPSRGTKEHLSITSTS 240  
191 ATQARPPPHAS-GPRRLG-----CESAMNHSVREAGVPLGLPAPGARRGGASGRS 231  
241 VPSAKKARCYPPVPRVERGP-----HRQVLPFSGKSN-VPSPARSPVPTAEKDLSSK 292  
232 LPLKPRRGAAPERTPVQGSWNAHPGTRGSDRGFCVVSFAR-----PAEATSLE 286  
293 GKVSGLSLG-SGVCCKKPKSTLSLSPRQNAQLRP-FIETHFLYSGDQGRNLNPSF 350  
287 GALSCTRHSFSGVGRHAGPSPSTSPRPPMDTPCPVPVVAETHFLYSSGD-KEQLRPSF 345  
351 LLSNLQPNLTGARLVEIIFLGSRPRTSGPLCRTHLSRRYWOMRPLFOQLLVNHAECY 410  
346 LLSLRPSLTGARLVEIIFLGSRPRTSGPLCRTHLSRRYWOMRPLFOQLLVNHAECY 405  
411 VLLRSHCRFTANQVTDAL-----NTSPFHLMDLLRLHSSPWQVY 452  
406 GVLLKTHCPLAA--VTPAAGVCAKPKQGSVAAPPEEDTPRLVQLLRQHSPPWQVY 462  
453 GFLPACILKVVASLWGTNRNRRFFKNLKKFISLGYKLSLOELMWKVEDCHNLS 512  
463 GFVACILRLVPPGLWGRNRRFFKNLKKFISLGYKLSLOELMWKVEDCHNLS 522  
513 SPGRDVPAAEHRILREILATFLWMDTVVQLLRSFYITESTFQKNRLFYKSVMS 572  
523 SPGVGCVPAEHRILREILAKFLHLMVSVVVELLSRFFVITETTFQKNRLFYKSVMS 582  
573 KLSIGVROHLERVLRELQOEVRHHDQTLWAMPICRLRFTPKPNGLRPIVNMYSMT 632  
583 KLSIGIRQHLKRVQLRELSEAEVRQREARPALLTSRLRFPKPDGLRPIVNMYSMT 642  
633 RALGRKQAOHFTORLTLFSLNLYERTKPHLMGSSVLGMNDIYETWFAFLVRLALDQ 692  
643 RTFPRKRAELTSRVKALSVNLNRYEARPGLGASVLGDDIHAWTFVLRVAQDP 702  
693 TPRMYFVKADVGTAYDAIPQKLVVEVVMNIRHSESTYCIQVAVVRDSQGVKHSFR 752

Db 703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIILK-PQNTYCVRRYAVVQKAAHGHVKAFKS 761  
Qy 753 QVTTLSDLOPYMGQFLKHQDSADALRNSVIEQISIMNESSSLDFDFLFLHSHSVK 812  
Db 762 HVSTLTDLOFYMRQFVAHLQET--SPLRDVAVIEQSSSLNEASGLFDFVFLRFMCHHAVR 819  
Qy 813 IGDRCTVQCGIPOGSSSLTLLCSLCFGDMENKLFABVORDGLLIRFVDDFLVTPHLDQ 872  
Db 820 IRKSYVQCGIPOGSSSLTLLCSLCFGDMENKLFABVORDGLLIRFVDDFLVTPHLDQ 879  
Qy 873 AKTFLTLVHGVPEYGCMLNOKTVNPFVPEPTLGGAAPIQLPAHCLFPMCGLLDQT 932  
Db 880 AKTFLTLVHGVPEYGCMLNOKTVNPFVPEPTLGGAAPIQLPAHCLFPMCGLLDQT 939  
Qy 933 LEVPCYSGVAOTSISITFQSVFKAGTKMENKLLSVLRKLCGSLFLDLQVNSLQTVCI 992  
Db 940 LEVQSYSSVARTSIRASLTFNRGFKAGNMRKLVGVLRKCHSLFLDLQVNSLQTVCI 999  
Qy 993 NIYKIFLQAYRFHACVIOQLPFDQVRKNLTFPLGIISQASCCYAILKVRNPGMTLKAS 1052  
Db 1000 NIYKILLQAYRFHACVIOQLPFDQVRKNLTFPLGIISQASCCYAILKVRNPGMTLKAS 1059  
Qy 1053 GS---FPPEAAHWLWQAEKLLAAHSVYKCLGLPLTAQKLLCKLPEATMTILKAAA 1109  
Db 1060 GAAGPLPSEAVOMLCHQAFLLKLTNRVTVYVPLLSLRTAQTSRKLPGTTLTAAEAAA 1119  
Qy 1110 DPALSTDFQITILD 1122  
Db 1120 NPALPSDFKILT 1132

RESULT 13  
AAE29226

ID AAE29226 standard; Protein; 1132 AA.

XX AAE29226;

XX 27-JAN-2003 (first entry)

XX Human telomerase reverse transcriptase (TERT).

XX Carbohydrate antigen; alpha(1,3)galactosyltransferase; alpha1,3GT; TERT;  
XX transgenic; alpha(1,2)fucosyltransferase; alpha1,2FT; human; enzyme;  
XX telomerase reverse transcriptase.

XX Homo sapiens.

XX WO200274948-A2.

XX 26-SEP-2002.

XX 21-MAR-2002; 2002WO-CA00378.

XX 21-MAR-2001; 2001US-277811P.

XX (GERO-) GERON CORP.

XX Denning C, Clark AJ, Schiiff JM;

XX WPI; 2002-759895/82.

XX N-PSDB; AAD46821.

XX Mammalian cells, useful for producing animal tissues with carbohydrate

XX antigens that are compatible for transplantation into human patients -

XX Disclosure; Page 34; 71pp; English.

XX The invention relates to animal tissues with carbohydrate antigens that  
XX are compatible for transplantation into human patients. The mammalian  
XX cell is inactivated homozygously for expression of alpha(1,3)galactosyl-  
XX transferase (alpha1,3GT) gene and comprises a transgene for alpha(1,2)-  
XX fucosyltransferase (alpha1,2FT). It is useful for producing animal tissue  
XX with carbohydrate antigens that are compatible for transplantation into



Sequence 1132 AA;

Query Match 59.4%; Score 3505; DB 23; Length 1132;  
Identical Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVARSLSRYREVPLATVRLGEGRELQVGPDPKIVRTIVACLVCEHW 60  
1 MTRAPRCVARSLSRYREVPLATVRLGEGRELQVGPDPKIVRTIVACLVCEHW 60  
61 GSQPPADLSFHQVSSLLKELVARVQVQLCERNENVLAFGELLNEARGGPPMAFTSSVR 120  
61 DARPPAAPSPQVSCLEKELVARVQVQLCERNENVLAFGELLNEARGGPPMAFTSSVR 120  
121 SYLNTVETLVRSGAMWLLSRVGDLLVYLLAHCAVLLVPPSCAYQVCGSPLYOICA 180  
121 SYLNTVETLVRSGAMWLLSRVGDLLVYLLAHCAVLLVPPSCAYQVCGSPLYOICA 180  
181 TTDWPSVSAVSRTRPVGRNTVRLFLQOIKSSROEAPKPLALPSRGTKRHLSTSTS 240  
181 ATQARPPPHAS - GPRRLG - - - - - CERAWNHSVREAGVPLGLPAPGARRRGGASRS 231  
241 VPSAKARCYVPRVRESGP - - - - - HRQVLPSPGKSW - VPSPARSPVPTREKDLSEK 292  
232 LPLPRPRGAAPERTPVQGGWNAHPGTRGSDRGFCVSPAR - - - - - PABEATSLE 286  
293 GKVSADLSL - GSVCCKHKPSLTLSPRONAFOLRP - FIETRHFLYSRDGOERLNPSF 350  
287 GALSOTRHSFSGVGRQHAGPSPSRFPDPWDTPCPVPVYAEATKHFYSSGD - KEQLRPSF 345  
351 LLSNLOPNTGARBVLII FLGSPRPTSGCLRTHRLSRVQWRLPFOOLLVNHAEQY 410  
346 LLSLRPSLTGARBLETIFIGSRPMPGTPRRLPLPQRYQWQRPFLIELLGHACQFY 405  
411 VRLLRSHCRFTANQOVTDAL - - - - - NTSPPHLMDLRLHSSPWQVY 452  
406 GVLLKTHCFLRAA - - - - - VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSPPQVY 462  
453 GFLRACLKTVSASLWTRHNERFPKMLKFTSLGKYGKLSLOELMWKMKVDDCHWLSR 512  
463 GFVACLRRLVPPGLGSRHNERFRLNTKFTSLGKHAKLSLOELTWKMSVRDCAWLR 522  
513 SPGRDVRPAASHRLRERLATAFLFELMDTVVQLLRFYFVITESTFOKRLFFYRKSVM 572  
523 SPGVGVCPAAEHLRERLAKLFLHLSVYVVELLRSFYFVITESTFOKRLFFYRKSVM 582  
573 KLSIGVQHLRVLRLSLEFVRHHQDTWLANPICLRPIPKNGRLPIVNMVSMGT 632  
583 KLSIGVQHLRVLRLSLEFVRHHQDTWLANPICLRPIPKNGRLPIVNMVSMGT 632  
633 RALGRKQAOHFTORLKTFLMNLVETKPHLMGMSVGLMNDIYRTWRAFLVRALDQ 692  
643 RTFREKGAERLTSEVKALFSLVNYEARPGLLGASVLDGDIHRAWRTFLVRQAQDP 702  
693 TPRMVFVADVTGAYDAI PQGLVEVVMIRHSESTYCTQVAVVRDQSGQVHKSPFR 752  
703 PPELYFVADVTGAYDTIPDRLTEVIAIK - PONTYCVRRYAVVQKAAGHVRKAFKS 761  
753 QVTLSDLPQVMGQFLKHLQSDASALRNSVVIQSI SMNESSSLSFDFFLHLFLRSVVK 812  
762 HVSTLTDLPYRQFVAHLQRT - - - - - SPLDVAVVIQSSLSNEASGLFDVFLRWCHAVR 819  
813 IGDCYTCQGI PQGSSILSTLCSLCFGDMENKLFABVQRDGLLLRFDVDFLVTPLHDQ 872  
820 IRGKYVQCGI PQGSSILSTLCSLCYGDENKLFAGIRDRDGLLLRFDVDFLVTPLHDQ 879  
873 AKTFLSTLVHGVPEYGCINLQKTVNFPFPGTLGGAAVQLPAHCLPWCGLLDTOT 932  
880 AKTFLTLVRGVPEYGCINLQKTVNFPFPGTLGGAAVQLPAHCLPWCGLLDTOT 939  
933 LEVFCDYSGYQTSIKTSLTFTQSVFKAGKTVRNKLSVLRLKCHGFLDLQVNSLQTVCI 992

QY 993 NIYKIFLLQAYRFHACVQLPFPQORVRKNLTFGLGISSOASCYAILKVNPGMTLKAS 1052  
Db 1000 NIYKIFLLQAYRFHACVQLPFPQORVRKNLTFGLGISSOASCYAILKVNPGMTLKAS 1059  
QY 1053 GS - - - - - FPPEAAHMLCYQAFLLKLAASHVIVKLLGPLRTAQKLLCKRLPEATWTILKAAA 1109  
Db 1060 GAAGPLPSEAVQWMLCHOAFILKILTRHRTVTVPLGLSLRTAQTLQSKLPFTTILEAAA 1119  
QY 1110 DPALSTDFQITLD 1122  
Db 1120 NPALPSDFKITLD 1132

RESULT 15  
ABP56676 standard; Protein; 1132 AA.  
XX ABP56676;  
XX AC  
XX DT 25-MAR-2003 (first entry)  
XX Human telomerase reverse transcriptase protein SEQ ID NO:2.  
XX Human; telomerase reverse transcriptase; enzyme; hTERT; chromosome 5;  
KW vulnary; antiulcer; epithelial cell migration promoter; wound;  
KW epithelial; skin wound; lesion; burn; surgical incision; ulcer;  
KW epithelial cell; keratinocyte; epidermal; mucosal.  
XX Homo sapiens.  
XX WO200291999-A2.  
XX PN 21-NOV-2002.  
XX PD 09-MAY-2002; 2002WO-US14867.  
XX PF 09-MAY-2001; 2001US-289903P.  
XX PR (GERO-) GERON CORP.  
XX PI Jiang X, Chiu C, Harley CB;  
XX WPI: 2003-120591/11.  
XX DR N-PSDB; ABZ22474.  
XX Composition for treating wounds and enhancing epithelialization of a skin surface, comprises vector encoding telomerase reverse transcriptase or telomerase epithelial cells on a microparticle or a matrix -  
PS Disclosure; Page 32; 68pp; English.  
XX The present invention describes a pharmaceutical composition (I) comprising a vector encoding telomerase reverse transcriptase (TERT) in an excipient or device, or comprises telomerase epithelial cells on a microparticle or a matrix suitable for topical administration or administration to a wound site. (I) has vulnary and antiulcer activities and can be used to promote epithelial cell migration. (I) is useful for treating a wound and enhancing epithelialization of a skin surface. The wound is especially skin wound including acute lesion such as traumatic lesion, burn, or surgical incision, chronic lesion such as chronic venous ulcer, diabetic ulcer or compression ulcer and the wound is further monitored for closure. The telomerase activity or TERT expression is increased in epithelial cells at the site of treatment. The also in fibroblasts or endothelial cells at the site of treatment. The epithelial cells are especially keratinocytes. A polynucleotide encoding TERT is useful for the preparation of a medicament for treatment of a wound or an epithelial surface in a human or animal. An epithelial cell with increased telomerase activity or increased expression of TERT is useful for preparation of a medicament for the treatment of a wound in a human or animal. (I) is also useful for treating wounds of other epidermal surfaces including mucosal surfaces such as bronchus, mouth,

nose, oesophagus, stomach, or intestine. The present sequence represents human TERT (hTERT), which is given in the exemplification of the present invention. hTERT is located to chromosome 5.

Sequence 1132 AA;

```
Query Match          59.4%; Score 3505; DB 24; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1  MTRAPRCVAVSLRSRVREVMPLATFVRRLGPEGRRLVQPGDPKIVRTIVAQCLVCMHW 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1  MPRAPRCVAVSLRSRVREVMPLATFVRRLGPEGRRLVQPGDPKIVRTIVAQCLVCMHW 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61  GSQPPADLSFHQVSSKELVARVVQBLCEENRVNLAFGFELLNARGPPMAFTSSVR 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
61  DARPPPAAPSFRQVSCIKELVARVQLCEGAKNVLAFGFALLDARGGPPPEAFTTSVR 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SYLPTNTVETLRVSGAMLLSRVGDLLVYLLAHCALVLLVPPSCAYOVCGSPLYQICA 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
121 SYLPTNTVDALRGSGWGLLRRVGGDVLVHLLARCALFVIVAPSCAYQVCGPPLYQLGA 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 TTDIWPSVASYRTPRVGRNFTNLRFLOQIKSSRQEPAPKPLALPSRGTKRHLSTSTS 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSASRS 231
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
241 VPSAKKACYPVPRVEGCP-----HQVLPTPSGKSW-VPSPARSPEVPTAEKDLSSK 292
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
232 LPLPKRPRGGAAPERTFVGQGSNAHPGTRGSPDRGFCVGPAP-----PAEATSLE 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
293 GKVDLSLS-GSVCKHKPSSLSLPPRQNAFQLRP-FIETRHFLYSRGDGOERLNPSF 350
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
287 GALSCTRSHSPSVGRQHAGPPSTSRPPMDTPCPVVAETKHFLYSSGD-KEQLRPSF 345
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
351 LLSNLQPNLTGARRLVEIFLGSRRPTSGPLCRTHRLSRRYWQWEPFLQOLLVNHAEQY 410
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
346 LLSLRPSLTGARRLVETIFLGSRRPMPGTFRRLPRLPQRYQWQRPFLLELLGNHAQCPY 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
411 VRLLRSHCRFRFANQQVTDAL-----NTSPHMLDILLRHSHPQVY 452
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
406 GVLKXTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSHPQVY 462
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
453 GFRLACLCKVVSASLWGRHNERFFQNLKFIISLGYKGLSLQELMWKVKVEDDCHWLSR 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
463 GFVRACLRLVPPGLWGRHNERFRLNKKFISLGHAKLSLQELTWKYSVRDCAWLR 522
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
513 SPQKRVPAAEHRLRERILATFLWMDTVYVQLLRSFFYITESTFQKNLFFYRKSVM 572
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
523 SPQVGCVPAAEHRLRERILAKELHLMVYVYVVELLRSSFFYVTFQKNLFFYRKSVM 582
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
573 KLQSIGVROHLRVLRELSEVEVRHQDTWLAMPICLRPIPKPNGLRPIVNMYSMTG 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
583 KLQSIGIRHLKRVQRELSEAEVRHQREARFALITSRLRPIPKPDGLRPIVNMYYVGA 642
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
633 RALGRKQKQHFQRLKTLFSLNLYERTKHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
643 RTFRREKRAERLTSRVKALFSLNTERARRPGLLGASVLGLDDIHRARWTFVLVRAQDP 702
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
693 TPRMYEVKADVTGAYDAIPQGLVEVWANNIRHSESTYCIQYAVVRDSDQGVHKSFR 752
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIHK-PONTYCVRRYAVQKAAHGHVRKAFKS 761
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
753 QVTLSDLPYMGQFLKHLQSDASALNSVVIQSI SMNESSSLPDPFHLRHSVVK 812
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 HVTSLTDLQFYMRQFVAHLQET--SPLRDVAVIEQSSLSNEASSGLFDVFLRFMCHAVR 819
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
813 IGDRCYTCQCGIPQGSLSLTLCSLCFCGDMENKLPFAEVRQDGLLIRFVDDFLVTPHLQ 872
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
820 IRGKSVYVQCGIPQGSILTLCSLCYCGDMENKLPAGIRRDGLLRLVDDFLVTPHLTH 879
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
873 AKTFLSTLVHGVPEYGCMLNQTVNPFVEPGTLGGAAPQLPAHCLFPWCGLLLDTQT 932
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
880 AKTFLRLVRGVPEYGCVMNLRKTVNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRI 939
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: February 4, 2004, 14:10:19

Job time : 51 secs

```
Qy 933 LEVFCYSGVTAQTSIKTSITFQSVFKAGTKMKNLLSVLRKCHGLFLDLQVNSLQTVCI 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 940 LEVQSDYSSYARTSIRASLTFRGFKAGNMRKLFGLRLKCHSLFLDLQVNSLQTVCT 999
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 993 NIYKIFLLQAYRFHACVIOLPFDQVRKXLTFFLGIISQASCCYAILKVNPGMTLKAS 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1000 NIYKIFLLQAYRFHACVIOLPFDQVRKXLTFFLGIISQASCCYAILKVNPGMTLKAS 1059
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1053 GS---PPPEAAHWLCYQAFLLKLAHSAVYIKCLLGLPLTAQKLLCRKLPKATMTILKAAA 1109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1060 GAAGPLPSEAVQWLCHQAFLLKLTTRHRYTVYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1110 DPALSTDFQTILD 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1120 NPALPSDFKILTD 1132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 4, 2004, 14:09:14 ; Search time 21 Seconds  
(without alignments)  
2260.609 Million cell updates/sec

le: US-09-042-460-2

ect score: 5901

ence: 1 MTRAPCPAVRSLRLRYRE.....TILKAADPALSTDFQITLD 1122

ing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

imum DB seq length: 0

imum DB seq length: 2000000000

:-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : Issued Patents:AA.\*

1: /cgn2\_6/prodata/1/iaa/5A-COMB.pep.\*

2: /cgn2\_6/prodata/1/iaa/5S-COMB.pep.\*

3: /cgn2\_6/prodata/1/iaa/6A-COMB.pep.\*

4: /cgn2\_6/prodata/1/iaa/6B-COMB.pep.\*

5: /cgn2\_6/prodata/1/iaa/6CTUS-COMB.pep.\*

6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

alt No.	Score	Query Match	Length	ID	Description
1	3505	59.4	1132	3	US-08-851-843A-225
2	3505	59.4	1132	3	US-08-974-549A-2
3	3505	59.4	1132	3	US-08-854-050-225
4	3505	59.4	1132	4	US-09-430-323-101
5	3505	59.4	1132	4	US-09-128-354-2
6	3505	59.4	1132	4	US-09-575-321-2
7	3505	59.4	1132	4	US-09-052-919-2
8	3505	59.4	1132	4	US-08-912-951-2
9	3505	59.4	1154	3	US-08-974-549A-611
10	3505	59.4	1154	4	US-08-912-951-323
11	3505	59.4	1189	3	US-08-974-549A-613
12	3505	59.4	1189	4	US-08-912-951-325
13	3505	59.4	1200	3	US-08-974-549A-612
14	3505	59.4	1200	4	US-08-912-951-324
15	3505	59.4	1285	3	US-08-974-549A-600
16	3505	59.4	1285	4	US-08-912-951-314
17	3496	59.2	1132	3	US-08-974-549A-344
18	3496	59.2	1407	3	US-08-974-549A-628
19	3496	59.2	1407	4	US-08-912-951-334
20	2492.5	42.2	1003	3	US-08-851-843A-217
21	2492.5	42.2	1003	3	US-08-974-549A-336
22	2492.5	42.2	1003	3	US-08-854-050-217
23	2492.5	42.2	1003	4	US-09-430-323-217
24	2195	37.2	807	3	US-08-974-549A-5
25	2195	37.2	807	4	US-08-912-951-5
26	1686.5	28.5	364	4	US-09-417-485D-41
27	1506.5	25.5	564	3	US-08-851-843A-101

Sequence 267, App  
Sequence 101, App  
Sequence 101, App  
Sequence 40, Appl  
Sequence 602, App  
Sequence 316, App  
Sequence 603, App  
Sequence 317, App  
Sequence 604, App  
Sequence 318, App  
Sequence 605, App  
Sequence 319, App  
Sequence 10, Appl  
Sequence 606, App  
Sequence 320, App  
Sequence 47, Appl  
Sequence 202, App

## ALIGNMENTS

### RESULT 1

US-08-851-843A-225  
; Sequence 225, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6093809e1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:



## SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

08-851-843A-225

very Match

est Local Similarity 59.4%; Score 3505; DB 3; Length 1132;

atches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCFAVASLRSRYREVMPLATVRLGPEGRRLVQGPDKIYRILVAQCIVCMHW 60

1 MPRAPRCFAVSLRSRYREVMPLATVRLGPEGRRLVQGPDKIYRILVAQCIVCMHW 60

61 GSOPPPADLSHQVSSIKELVARVQBLCEBENRNLAFGELLNEARGPPMATTSVR 120

61 DARPPAPSPFQVSCLEKELVARVQBLCEBENRNLAFGELLNEARGPPMATTSVR 120

121 SYLPTVITLTVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYOVCGSPLYQICA 180

121 SYLPTVITLTVSGAWMLLSRVGDDLLVLLAHCALVLLVPPSCAYOVCGSPLYQICA 180

181 TTDIWPVSASYPTRPVGRNFTNLRFLOQIKSSROEAPKPLALPSRGTKHLSLTSTS 240

181 ATQARPPPHAS-GPERRRLG-----CERAWNSVREAGVPLGLPAPGARRRGSSASRS 231

241 VPSAKKACYPVPRVEGCP-----HRQVLPSPGKSW-VPSPARSPVPTAEKDLSSK 292

232 LPLPKRPRGAAPERTFVQGSWNAHPGTRGSPDRGFCVVSFAR-----PAEATLSLE 286

293 GKVDLSLS-GSVCCKKXPSSTLSPPRQNAFQLRP-FIETRHFLYSRGDQERLNPSF 350

287 GALSCTRSHPSVGRQHEAGPSTSPRPMDTPCPVYAEATKHFLYSSGD-KEQLRPSF 345

351 LLSNLOPNLTGARLVELIFLGSPPRSGPLCRTHLSRYWQWRLPQOLLVNHAEQY 410

346 LLSLRPLTGARLVELIFLGSPPRSGPLCRTHLSRYWQWRLPQOLLVNHAEQY 405

411 VLLASHCRFRFTANQQVTDAL-----NTSPPHLMDLLRLHSSPWQVY 452

406 GVLLKTHCPRAA---VTPAAGVCAKPKQGSVAAPPEEDTPRRLVQLLRQSSPWQVY 462

453 GFLRACLCKVNSASLWTRNRRERFFKNLKKFISLKGKGLSLQELMWKVKVDDCHWLS 512

463 GFVRACLRLVPPGLWGRSHRERFLRNTKFKISLGHAKLSLQELTWKNSVEDCAWLR 522

513 SPCKDRVPAASHRLRERILATFLWMDTVYVQLLRFFVITESTFOKRLFFYKSVMS 572

523 SPGVGCVPAASHRLRERILATFLWMDTVYVQLLRFFVITESTFOKRLFFYKSVMS 582

573 KLOSIGVQRHLERVLRELRSQEVRRHQDTWLMPICLRPIPKPGLRPIVNSYSMTG 632

583 KLOSIGIRHLKRVQLRELSEAEVQRHREARFALLTSRLRFIPKPDGLRPIVNMVYVGA 642

633 BALGRKQAOHTORLKLTPSLMANYRTKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692

643 RTFREKRAELTSRVKALFVNLVNERARRCLLGASVLGDDIHRARWTFVLRVRAQDP 702

693 TPRMYFVADVTGAYDAIPQGLKVEVVMNIRHSESTYCIQVAVVRDSDQGVKHSFR 752

703 PPELYFVKVDVTGAYDTIPQDLTEVIAIHK-PQNTVCVRRVAVVQAAHGHVRKAFKS 761

753 QVTTLSDLQPYMGFLKHLQSDASANSVVEQSI SMNNESSSLDFDLHFLHSVVK 812

762 HVSTLTLDQPYNRQVVAHQET--SPLSDAVVIEQSSSLNEASSGLFDFLRMCHAVR 819

813 IGDRCVTCQCGIPQGSSTLSLCSLCFQDMENKLFVQVQDGLLRVDDFLVLTPHLQ 872

820 IRKSVQVQCGIPQGSSTLSLCSLCFQDMENKLFVQVQDGLLRVDDFLVLTPHLQ 879

873 AKTFSLTLVHGVPEYGCMMINOKTVNPNFPGTLGGAAVQPLPAHCLFPWCGLLDTQT 932

Db 880 AKTFSLTLVHGVPEYGCMMINOKTVNPNFPGTLGGAAVQPLPAHCLFPWCGLLDTQT 939

Qy 933 LEVFCDSYGAQTSIKTSLTFSVFKAGTKWRNKLKSLVLRKCHGLFLDLQVNSLQTVCI 992

Db 940 LEVQSDYSYARTSIRASLTFRNGFKAGNRKRLFGVLRKCHSLFLDLQVNSLQTVCT 999

Qy 993 NIYKIFLLQAYRFHACVQLPFDORVRNLTFFLGISSQASCCVAILKVNPGWTLKAS 1052

Db 1000 NIYKILLQAYRFHACVQLPFPHQVQWKNPFFFLFVDSITASLSCYILKAKNAGWSLGAK 1059

Qy 1053 GS---PPPEAAHWLCYQAFLLKLAHSHSVIYKLGPLRTAQKLLCRKLPEATWTLKAAA 1109

Db 1060 GAAGPLPSEAVQWLCHQAFLLKLTNRHRTVYVFLGSLRTAQQLSRKLPGLTTLTALEAAA 1119

Qy 1110 DPALSTDFOTILD 1122

Db 1120 NPALPSDFXTILD 1132

## RESULT 2

US-08-974-549A-2

; Sequence 2, Application US/08974549A

; Patent No. 6186178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Haxley, Calvin B. H.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/974,549A

; FILING DATE: 19-NOV-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/915,503

; FILING DATE: 14-AUG-1997

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

## ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-0026-00US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO. 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

-08-974-549A-2

Query Match 59.4%; Score 3505; DB 3; Length 1132;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCRAVRLLRSRYEVNPLATFVRRLLGPEGRRLVQPDGPKIVRTLVAAOCLVCMHW 60  
1 MPRAPRCRAVRLLRSRYEVNPLATFVRRLLGPEGRRLVQPDGPKIVRTLVAAOCLVCMHW 60  
61 GSQPPADLSHFQVSSKELVARVQRCERNERNVLAFFGELLNEARGPPMAFTSSVR 120  
61 DARPPAPAPFRQVSCLELVARVQLRCERGAKNVLAFFGELLNEARGPPMAFTSSVR 120  
121 SYLNPVTIETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPSCAVQVCGSPLOYICA 180  
121 SYLNPVTIETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPSCAVQVCGSPLOYICA 180  
181 TTDIWPVSASVYRTPVGRNFTNLRFQIIXSSSQEAPKPLPALPSGTRKRLSLTSTS 240  
181 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAPGARRRGGASRS 231  
241 VPSAKKACVYPRVVEEP-----HQVLPTRSGKSW-VPSAPSPVEPTAEKOLSSK 292  
232 LPLKRRRGAAPERTFVGGSWAHGPRTRGSDGFCVVSAPR-----PAEATSLE 286  
293 GKVSLSLS-GSVCKKHPSSLSLSPRQNAFQLRP-FIETRHLYSRGDSGQERLNPSP 350  
287 GALSCTRSHSPSVGQHAGPPSTSRPRPNDTPCPVYAEKHFLYSSGD-KEQLPSP 345  
351 LLSNLPNTGARRLVEIFLGSRRPTSGPLCRHRLSRRYQWRPLFOQLLVNHAECQY 410  
346 LLSLRPSITGARRLVEIFLGSRRPTSGPLCRHRLSRRYQWRPLFOQLLVNHAECQY 405  
411 VLLRSHCRFRNTANQVTDAL-----NTSPHMLDLRLHSSPMQVY 452  
406 GVLLKTHCPLRAA---VTPAGVCAREKYPQSVAAPEEDTPRRVLQLRQSSPMQVY 462  
453 GFLRACLKVVASLWGRTHNRERRFKNLKFIISLGYKGLSLQELMWMKVEDCHWLS 512  
463 GFVRACLRLVPPGLWGRTHNRERRFKNLKFIISLGYKGLSLQELMWMKVEDCHWLS 522  
513 SPKORVPAEHLRERILATFLWMTYVOLLRSFFYITESTFQKRLFFYKRSVMS 572  
523 SPVGVCPAAEHLRERILATFLWMTYVOLLRSFFYITESTFQKRLFFYKRSVMS 582  
573 KLSQIGVGRHLRERILRSLQSEVHHQDTHLAMPICRLRPIKPNGLRPIVNMYSMTG 632  
583 KLSQIGVGRHLRERILRSLQSEVHHQDTHLAMPICRLRPIKPNGLRPIVNMYSMTG 642  
633 RALGRKQACHTQRLKTLFSLNRYERKHPHMGSSVLGMNDIYRTWRFAVLRLALDQ 692  
643 RTFRREKRAELTSLRVKALFSLVNLRYERARREPLIGASVLGLDDIHRAWRTFVLVRADP 702

## RESULT 3

US-08-854-050-225

; Sequence 225, Application US/08854050

; Patent No. 6261836

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: No. 6261836el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/854,050

; FILING DATE: 09-MAY-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00293005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 225:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

i-08-854-050-225

Query Match

Best Local Similarity 59.4%; Score 3505; DB 3; Length 1132;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVRSLLRSYREVMPLATFVRRLGPEGRRLVQPGDKIYRLVAQCILVCMHW 60

1 MPRAPCRAVRSLLRSYREVMPLATFVRRLGPEGRRLVQPGDKIYRLVAQCILVCMHW 60

61 GSQPPADLSFHQVSSKELVARVQRLCERNRNVLAFGFELLNEARGPPMAFTSVR 120

61 DAREPPAAPSPFRQVSCIKELVARVQRLCERNRNVLAFGFELLNEARGPPMAFTSVR 120

121 SYLPTNVTIELRVSGAWMLLSVGGDLLVYLLAHCALVLLVPBSCAVCGSPLYQICA 180

121 SYLPTNVTIELRVSGAWMLLSVGGDLLVYLLAHCALVLLVPBSCAVCGSPLYQICA 180

181 TTIDWPVSASVYRPTVRVGRNFTNLRFLOQIKSSSRQEAPEKPLALPSRGTGRHLSLTSS 240

181 ATQARPPPHAS-GPFRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS 231

241 VPSAKKARCVPVRVEBP-----HRQVLPSPGKSM-VSPARSPEVPAEKDLSK 292

232 LPUPKPRFGAEPERTPVQSGSWAHQKTRGSDRGFCVVSAPAR-----PAEATSLE 286

293 GKVDLSLS-GSVCKHKPSTSLSPRONAFQLRP-FIETRHFLYSGDQGERLNPSF 350

287 GALSGRHSPSVGRQHAGCPSTSRPPRWDTPCPVVAETKFLYSSGD-KEQLRPSF 345

351 LLSNLQNLGARRLVEIIFLGRSPTSGLCKTHLSRRYQWMPPLFOQLLVNHAQCY 410

346 LLSLRPSLTGARRLVETIFLGRSMPGTPRLPQLPQRYQWMPPLFLELLGNHAQCPY 405

411 VRLRSRCHRTANQOVTDAL-----NTSPHMLDILLRHSSPQWVY 452

406 GVLLKTHCPILRAA---VTPAAGVCAREKPGQSVAAPEEDTPRLVLQRLROHSSPQWVY 462

453 GFLRACLCKVVSASLWTRNRRFPNLIKFTSLGKYKLSLQELMMKMKVEDCHWLS 512

463 GFVRACLRRLVPPGLWGSRRNRRFLNTKFTSLGKHAKLSLQELTWKMSVRDCAWLR 522

513 SPGKDRVPAEHLREILATELFLWMDTVVOLLRSFFVITESTFQKRLFFYKSVMS 572

523 SPGVGCVPAEHLREILAKFLHMSVTVVLLRSFFVITESTFQKRLFFYKSVMS 582

573 KLSIGVROHLRVRRLRSQEBVRHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMTG 632

583 KLSIGIRQHLKRVQLRELSEAEVRQREARPAALLSRFLRFPKPDGLRPIVNMVYVGA 642

633 RALGRKQAQHOFTQRLKTLISMLNRYETKPHLMGSSVLGMDIYRTWEAFVLRVRLDQ 692

643 RTFRREKRAELTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRAMFTFVLRVRAQDP 702

QY 693 TPRMTFVADVTGAYDAIPQKLVVEVNVANMIRHSESTYCIROYAVVRDSQGVHKSFR 752  
DB 703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIHK-PQNTYCVRRYAVVQKAHGHVRKAFKS 761  
QY 753 QVTTLSDLQPYMGQFLKHLQSDASALRNSVVIQSSISMNESSSSLPDDFFLHFLHRSVVK 812  
DB 762 HVSTLTDLQPYRQFVAHQET--SPLEDAVVIEQSSSLNEASSGLFDVFLRFMCHAVR 819  
QY 813 IGDRYTCQCGIPQSSSLSTLLCSLCFGDMENKLFARVORDGLLLRRFVDDFLVTPHLQ 872  
DB 820 IRKGSYVQCGIPQSSSLSTLLCSLCYGDENKLFAGIRRDGLLLRLVDDFLVTPHLTH 879  
QY 873 AKTELSTLVHGVPEYGCINLQKTVNFPVBPGLGGAAPYQLPAHCLFWCGLLLDTQT 932  
DB 880 AKTFLRTLVRGVPEYGCINLQKTVNFPVBPGLGGAAPYQLPAHCLFWCGLLLDTRT 939  
QY 933 LEVPCDYSYATSIKTSLSLTSQSVFKAGKTVNKLVLRLKCHGLFLDLQVNSLQVCI 992  
DB 940 LEVQDYSYATSIKTSLSLTSQSVFKAGKTVNKLVLRLKCHGLFLDLQVNSLQVCI 999  
QY 993 NYIKFLLQAYRFAHCVIQLPDRQVRKNLTFELGIISSQASCCYAILKVKNGMTLKAS 1052  
DB 1000 NYIKFLLQAYRFAHCVIQLPDRQVRKNLTFELGIISSQASCCYAILKVKNGMTLKAS 1059  
QY 1053 GS---FPPEAAHNLQYAFLLKLAHSHVYKLLGLPLRTAQKLCRLKPEATMTILKAAA 1109  
DB 1060 GAAGPLPSEAVQWLCHQAFLLKLTETRVYVYVPLLSLRTAQKLCRLKPEATMTILKAAA 1119  
QY 1110 DPALSTDFTILD 1122  
DB 1120 NPALPSDFKTLID 1132

## RESULT 4

US-09-430-323-225

; Sequence 225, Application US/09430323

; Patent No. 6309867

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; Lingner, Joachim

; Nakamura, Toru

; Chapman, Karen B.

; Morin, Gregg B.

; Harley, Calvin

; Andrews, William H.

; TITLE OF INVENTION: No. 6309867el Telomerase

; NUMBER OF SEQUENCES: 225

; CORRESPONDENCE ADDRESS:

; ADDRESS: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/430,323

; FILING DATE: 29-Oct-1999

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-0029300S  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 225:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
 19-430-323-225

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSLLRSRYREVNPLATFVRRLLGPEGRLVQPGDPIKIVRTLVAAQCLVCMHW 60  
 1 MPAPRCRAVRSLLRSRYREVNPLATFVRRLLGPEGRLVQPGDPAAPALVAQCLVCPWF 60  
 61 GSOPPPADLSFQVSSKLKELVARVQRCERNERNVLAFGPELLNEARGGPPMAFTSSVR 120  
 61 DARPPPAAPSFQVSCLEKELVARVQLQRCERGAKNVLAFGFALLDGAAGGPPPEAFTTSVR 120  
 121 SYLPNTVETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYOICA 180  
 121 SYLPNTVETLRVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYOICA 180  
 181 TTDWPSVGSASVPTPVGNFTNLRFQIQQSSRQEPKPLALPSPRGTKRHLSTLS 240  
 181 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPADGARRRGCSASRS 231  
 241 VPSAKKARCVPRVEEGP-----HRQVLTTPSGKSW-VPSPARSEVTAEDKLSK 292  
 232 LPLPKPRGAAPERTPVGQGSWAHPGRTGSDRGFCVVSPPAR-----PABEATSL 286  
 293 GKVSIDLSS-GSVCCCHKFSSTLSLPPQNAQLRP-PIETRHPLYSRGQDQLNAPSF 350  
 287 GALSGRTHSHPSVGRQHAGPSTSRPPRPDTPCPVVAETKHFLYSSGD-KQLRPSF 345  
 351 LLSNLOPLTGARLVEIIFLGSRPRTSGPLCRTHLSRRTWQMPLEFQQLLVNHAQQY 410  
 346 LLSLSPLSTGARLVETIFLGSRPMPGTPRRLPQLPQRYWQMEPLFELLGNAQQY 405  
 411 VLLRSHCRPRTANQOVTDAL-----NTSPPHMLDLRLHSSPQVY 452  
 406 GVLLKTHCPILRAA---VTPAAGVCAREKPGQSVAAPEBEDTDFRLVOLLQHSPPQVY 462  
 453 GFIRACLVKVSASLMTGRNRRFPKLNKFTISLGKYGKLSQELMWMKVBCHMLRS 512  
 463 GFVRACLRVLPPVLMGSRHNRFRNLTXTISLGKIAKLSQELTWKMSVRDCAMLR 522  
 513 SPKDRVPAEHLREIILATFLWMDTVVQLARSFYLTSTFQKNRLFVFRKSVWS 572  
 523 SPVGVCVPAEHLREIILAKFLHMLNSVYVELLRSFYVTEITFQKNRLFVFRKSVWS 582  
 573 KLSIGVQRLERVLRLSLSQEVHRHQDTWLAMPICRLRIFPKNGRLPIVNMVSMGT 632  
 583 KLSIGVQRLERVLRLSLSQEVHRHQDTWLAMPICRLRIFPKNGRLPIVNMVSMGT 642  
 533 RALGRKQAGHPTQRLKTLFSLMYERTKPHLMGSSVLGMNDIYRTWRVILVRLDQ 692  
 643 RTRFRKRAERLTSRVKALFSLVINYERARPGGLGASVLGDDIHRWRTFVLVRAQDP 702  
 693 TPRIYFKADVDTGAYDAIQGKLVEVANNIRHSSTYCIQYAVVRDSDQGVHKSPFR 752  
 703 PPFLYFKVDVDTGAYDTIQDRLTEVIAIIR-PQNTYCVRYAVVQKAAGHVRKAFKS 761

QY 753 QVTTSLDLPQYMGOLFKEHLOSDASALRNSVWIEQISIMNESSSLFDFELHFLRHSVVK 812  
 DB 762 HVSUTLDLPQYMRQFVAHQET--SPLRDAVWIEQSSLINEASSGLFDVFLRPMCHAVR 819  
 QY 813 TGDRCYTCQCIPOGSSSLTLLCSLCFGDMENKLFARVQRDGILLRPFVDDFLVTPHLDQ 872  
 DB 820 TRGKSYVQCQIPOGSSILTLLCSLCYCDGMENKLFAGIRDRDGLLRVDDFLVTPHLDQ 879  
 QY 873 AKTFLSTLVHGVPEYGVGMINLOKTVVNFPPVPGTLGGAAPVQLPAHCLPFWCGLLDTOT 932  
 DB 880 AKTFLSTLVHGVPEYGVGMINLOKTVVNFPPVPGTLGGAAPVQLPAHCLPFWCGLLDTOT 939  
 QY 933 LEVFCDSYGYAQTISKTSLTFQSVFKAGTKMKNKLSVLRLKCHGLFLDLQVNSLQTVCI 982  
 DB 940 LEVQSDYSYARTSIRASLTFRNGFKAGRNMRKLFGLVRLKCHSLFLDLQVNSLQTVCT 999  
 QY 993 NYTKIFLQAVRFHACVLTQPPDQVREKNLTFFELGIISSOASCYAILKLVNPGMTLKAS 1052  
 DB 1000 NYTKIFLQAVRFHACVLTQPPDQVREKNLTFFELGIISSOASCYAILKLVNPGMTLKAS 1059  
 QY 1053 GS--FPPEAAHMLCYQAFLLKLAHSHVIYKCLIGPLRTAQKLLCRKLPKPEATMTILKAAA 1109  
 DB 1060 GAAGPLPSEAYQMLCHQAFLLKTRHRVTVYVPLLGSLRTAQKLLCRKLPKPEATMTILKAAA 1119  
 QY 1110 DPALSTDFQTILD 1122  
 DB 1120 NFALPSDFKTILD 1132

RESULT 5  
 US-09-128-354-2  
 ; Sequence 2, Application US/09128354  
 ; Patent No. 6337200  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morin, Gregg B.  
 ; APPLICANT: Geron Corporation  
 ; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants  
 ; FILE REFERENCE: 015389-003310US  
 ; CURRENT APPLICATION NUMBER: US/09/128,354  
 ; CURRENT FILING DATE: 1998-08-03  
 ; EARLIER APPLICATION NUMBER: US 08/851,843  
 ; EARLIER FILING DATE: 1997-05-06  
 ; EARLIER APPLICATION NUMBER: US 08/854,050  
 ; EARLIER FILING DATE: 1997-05-09  
 ; EARLIER APPLICATION NUMBER: US 08/911,312  
 ; EARLIER FILING DATE: 1997-08-14  
 ; EARLIER APPLICATION NUMBER: US 08/912,951  
 ; EARLIER FILING DATE: 1997-08-14  
 ; EARLIER APPLICATION NUMBER: US 08/915,503  
 ; EARLIER FILING DATE: 1997-08-14  
 ; EARLIER APPLICATION NUMBER: WO PCT/US97/17618  
 ; EARLIER FILING DATE: 1997-10-01  
 ; EARLIER APPLICATION NUMBER: WO PCT/US97/17885  
 ; EARLIER FILING DATE: 1997-10-01  
 ; EARLIER APPLICATION NUMBER: US 08/974,549  
 ; EARLIER FILING DATE: 1997-11-19  
 ; EARLIER APPLICATION NUMBER: US 08/974,584  
 ; EARLIER FILING DATE: 1997-11-19  
 ; EARLIER APPLICATION NUMBER: US 09/052,864  
 ; EARLIER FILING DATE: 1998-03-31  
 ; NUMBER OF SEQ ID NOS: 21  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1132  
 ; TYPE: PPT  
 ; ORGANISM: Homo sapiens  
 ; US-09-128-354-2

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPCPAVRSLLRSRYREVNPLATFVRRLLGPEGRLVQPGDPIKIVRTLVAAQCLVCMHW 60

```
1110 DPALSTDFQITLD 1122
1120 NPALPSDFKITLD 1132

RESULT 6
US-09-675-321-2
; Sequence 2, Application US/09675321
; Patent NO. 6440735
; GENERAL INFORMATION:
; APPLICANT: Gaeta, Federico C.A.
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune
; TITLE OF INVENTION: Response to a Telomerase Antigen
; FILE REFERENCES: 015389-003500FC
; CURRENT APPLICATION NUMBER: US/09/675,321
; CURRENT FILING DATE: 2000-09-28
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-675-321-2

Query Match 59.4%; Score 3505; DB 4; Length 1132;
Best Local Similarity 62.4%; Pred. No. 0;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCVRSLLRSRYREVWPLATFVRRLGPEGRLVQDGPDKIYRVLVAQCLVCMHW 60
DB 1 MPAPRCVRSLLRSRYREVWPLATFVRRLGPEGRLVQDGPDKIYRVLVAQCLVCMHW 60
QY 61 GSQPPADLSFHQVSSKLKELVARVQRLCERNERNVLAQFELLNEARGPPMAFTSSVR 120
DB 1 MPAPRCVRSLLRSRYREVWPLATFVRRLGPEGRLVQDGPDKIYRVLVAQCLVCMHW 60
QY 61 DARPPAPAFRSQVCLKELVARVQRLCERNERNVLAQFELLNEARGPPMAFTSSVR 120
DB 1 MPAPRCVRSLLRSRYREVWPLATFVRRLGPEGRLVQDGPDKIYRVLVAQCLVCMHW 120
QY 121 SYLNTVETLRVSGAMWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
DB 1 SYLNTVETLRVSGAMWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
QY 121 SYLNTVETLRVSGAMWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
DB 1 SYLNTVETLRVSGAMWLLSRVGDLLVYLLAHCALVLLVPSCAYQVCGSPLYQICA 180
QY 181 TTDIWPVSASVYRTPRVGRNFTNLRFLQIKSSRQEAPEKPLALPSRGTKRHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS 231
QY 241 VPSAKKARCYPVPRVEGP-----HRQVLPSPGSKW-VSPARSPEVPTAEKDLISK 292
DB 241 VPSAKKARCYPVPRVEGP-----HRQVLPSPGSKW-VSPARSPEVPTAEKDLISK 292
QY 232 LPLPRPRRGAPEPERTFVGQSWAHQPGTRGSDRGFCVSPAR-----PABEATSLE 286
DB 232 LPLPRPRRGAPEPERTFVGQSWAHQPGTRGSDRGFCVSPAR-----PABEATSLE 286
QY 293 GKVDLSLS-GSVCKKHPSTSLSPRQNAFQLRP-FIETRHFLYRGDQOERLNSPF 350
DB 293 GKVDLSLS-GSVCKKHPSTSLSPRQNAFQLRP-FIETRHFLYRGDQOERLNSPF 350
QY 287 GALSCTRSHSPSVGRQHAGPSTSRPPWDTPCPVYAEYKHFYSSGD-KEQLRPSF 345
DB 287 GALSCTRSHSPSVGRQHAGPSTSRPPWDTPCPVYAEYKHFYSSGD-KEQLRPSF 345
QY 351 LLSNLQNLGARRLVEIFLGSRRPTSGPLCRTHLSRRYQWQMRPLFOQLLVNHAECY 410
DB 351 LLSNLQNLGARRLVEIFLGSRRPTSGPLCRTHLSRRYQWQMRPLFOQLLVNHAECY 410
QY 346 LLSLRPSLTGARRLVEIFLGSRRPTSGPLCRTHLSRRYQWQMRPLFOQLLVNHAECY 405
DB 346 LLSLRPSLTGARRLVEIFLGSRRPTSGPLCRTHLSRRYQWQMRPLFOQLLVNHAECY 405
QY 411 VLLRSRCHRTANQQVTDAL-----NTSPPHLMDLLRHSFPQVY 452
DB 411 VLLRSRCHRTANQQVTDAL-----NTSPPHLMDLLRHSFPQVY 452
QY 406 GVLKTHCPHRAA---VTPAGVCAREKPGQSVAAPEEDTDPRLVOLLRQSSFPQVY 462
DB 406 GVLKTHCPHRAA---VTPAGVCAREKPGQSVAAPEEDTDPRLVOLLRQSSFPQVY 462
QY 453 GFLRACLCKVSVASLWGRTRHNERFFKNLKFISLGKYLKSLQELMKNKVEDCHWLS 512
DB 453 GFLRACLCKVSVASLWGRTRHNERFFKNLKFISLGKYLKSLQELMKNKVEDCHWLS 512
QY 463 GFVRACLRLVPPGLWGRSHNERFFKNLKFISLGKYLKSLQELMKNKVEDCHWLS 522
DB 463 GFVRACLRLVPPGLWGRSHNERFFKNLKFISLGKYLKSLQELMKNKVEDCHWLS 522
QY 513 SPGRDVPAAEHLRERILATFLFWLMDTVVOLLRSFFVITESTFQKRLFFYKSVMS 572
DB 513 SPGRDVPAAEHLRERILATFLFWLMDTVVOLLRSFFVITESTFQKRLFFYKSVMS 572
QY 513 SPGRDVPAAEHLRERILATFLFWLMDTVVOLLRSFFVITESTFQKRLFFYKSVMS 572
DB 513 SPGRDVPAAEHLRERILATFLFWLMDTVVOLLRSFFVITESTFQKRLFFYKSVMS 572
```

523 SPGVGCVPAEHRLEBILAKLHMLMSVTVVWELLRSFFVVTETTFQKNRLFFYRKSVMS 582  
573 KLOSIGVROHLERVRRLSQQEEVRRHQDWTWLPICRLRFIPKPNGLRPIVNMYSYMG 632  
583 KLOSIGVROHLKRVQLRSEAEVRQHRARPALLTSRLRFIPKPDGLRPIVMDYVYGA 642  
633 RALGRKQAOHFTORLTKLSMLNYRTKPHLMGSSVLGMDIYRTWPAFVLRVRLDQ 692  
643 RTFERREKRAELTSRVKALPSVLNYEARPPGLGASVLGDDIHRARWTFVLRVRAQDP 702  
693 TPRMYFVKADVGTAYDAIPQGLVEVVMNIRSESTYICIROVAVVRDSSQOVHKSFR 752  
703 PPELYFVKVDVTGAYDTIPQDLRTVEIASIK.PQNTYCVRRYAVVQKAAGHVRKAKS 761  
753 QVTTLSLOPYMGQFLKHLQSDASALRNSVWIEQISWNSESSSLFDFLHFLRHSVVK 812  
762 HVSTLTDLQPMQFVAHQET--SPLRDVAVIEQSSSLNEASSGLFVFLRFMCHHAVR 819  
813 IGRDCTQCQGIPOGSSLSLTLCSLQFGDMENKLFABVORDGILLRFDVDFLLVTPHLDQ 872  
820 IRGKSYVQCQGIPOGSSLSLTLCSLQFGDMENKLFAGIRDRDGLLLRLVDVFLVTPH 879  
873 AKTFLSTLVHGVPEYGCWMLNQLKTVNFPVEPGTLCGAAPYQLPAHCLFPWCGLLDQT 932  
880 AKTFLTLVRGVPEYGCWMLNQLKTVNFPVEDEALGTAFAVQMPAHGLFPWCGLLDRT 939  
933 LEVCDYSGVAQTSIKTSLTFSQSVKAGTMRNKLJSLVRLKCHGLFLDLQVNSLQTVCI 992  
940 LEVQSDYSSVARTSIRASLTFRNGKAGRNMRKLFGLRLKCHSLFLDLQVNSLQTVCT 999  
993 NIYKIFLLQAYRFHACVQIQLPQVRKNTLFFLGIISSQASCCYAILKVKXNPGMTLKAS 1052  
1000 NIYKILLQAYRFHACVQIQLPQVQWKNFTFLRVISDTASLCYSILKAKNAGSLGAK 1059  
1053 GS----FPEBAHWLCYQAFLLKLAHNSVIYKLLGPRTAQKLLCRKLPERNTWILKAAA 1109  
1060 GAAGLPSEAVQWLCYQAFLLKTRHRTVYVPLLSGLRTAQTLQSLKPLGTTTLTLEAAA 1119  
1110 DPALSTDFTQILD 1122  
1120 NPAFPSDFKILD 1132

ULT 7

09-052-919-2

sequence 2, Application US/09052919

atent No. 644650

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.

APPLICANT: Lingner, Joachim

APPLICANT: Nakamura, Toru

APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.

APPLICANT: Harley, Calvin B.

APPLICANT: Andrews, William H.

TITLE OF INVENTION: Antisense Compositions for Detecting and

TITLE OF INVENTION: Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/052,919

FILING DATE: 31-MAR-1998

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,549  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/974,584  
FILING DATE: 19-NOV-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Parent, Annette S.  
REGISTRATION NUMBER: 42,058  
REFERENCE/DOCKET NUMBER: 015389-003600US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-052-919-2

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

Qy 1 MTPAPRCRAVRSLLRSRYREVWPLATFVRIGCGERRLVQCDPKIYRTLVQAQCLVCMHW 60  
Db 1 MPRAPRCRAVRSLLRSRYREVLPLATFVRRLGPGQWRLVQRGDPAAFALVAQCLVCPW 60  
Qy 61 GSOPPPADLSFHQVSSSLKELVARVVQRLCERNERNVLAFFGPELLNEASGPPMAFTSSVR 120  
Db 61 DARPPAAPSPFQVSCSKELVARVQLRCERAKNVLAFFGALLDARGGPPFAFTTSVR 120  
Qy 121 SYLPTNTVIETLRVSGAWMLLSRVGDDLLVYLLAHALCALYLLVPPSCAYQVCGSPLYQCA 180  
Db 121 SYLPTNTVDALRGSGAWGLLRVVGDDVLVHLARCALFVLVAPSCAYQVCGPPPLYQLGA 180  
Qy 181 TTDIWFSVASYPPTFPVGNFTNLRFQOIKSSSRQEPKPLALPSGTGKHLSTLSTS 240  
Db 181 ATQARPPPHAS-GPRRLG-----CERANWHSVRGAVGLPLAPGARRRGSSASRS 231  
Qy 241 VPSAKKACVYPVPEVEGP-----HROVLTPSPKSGW-VPSPARSPVPTAEKDLSSK 292

```

222 LPLPKRRRGAAPBETPTVQGSNAHPGTRGSPGCVVSPAR-----PAEATSUE 286
293 GKVDLSLS-GSVCKKPKSTSLSPRONAQLRP-FIETRHFLYSRGQOERLNFSF 350
287 GALSGRHSHPVGRQHAGPSPSTRPRPDWTPCPVVAETKHFLYSSGD-KEQLRPSF 345
351 LLSNLOPNLTGARBVEIIFLGSRPRTSGPLRTHLSRYWCMRPLFOQLLVNAECQY 410
346 LLSLRPSLTGARBVEIIFLGSRPMPGPRRLPRLPQRYQWRPLFLELLGNACQPY 405
411 VLLRSHCRFTANQOVTDAI-----NTSPPHLMDLLRLHSSPWQVY 452
406 GVLLKTHCPLRAA--VTPAAGVCAREKPGSVAAPBEEEDTPRLVQLLRHSSPWQVY 462
453 GFLRACLCKVVSASLWTRNRRNRRFFKNLKFISLGYKGLSLQELMWMKVEDCHWLS 512
463 GFVRACURRLVPPGLWGRSRNRRFFLENTKFFISLGHAKLSLOELTWKQSVRDCAWLR 522
513 SPGKDRVPAAEHRLRERILATFLFELMDTVVQLLRSSFFYITESTFQKRLFFYKSVMS 572
523 SPGVCVPAAEHRLRERILAKFLHLMVSVVVELLRSSFFYVTTTQKRLFFYKSVMS 582
573 KLSQIGVQHLEVRVRLRELSOEVRHODTWLAMPICRLRFIPKPNGLRPIVMSYSGT 632
583 KLSQIGIRQLKRVQLRELSAEVROHREARPAALLTSRLRFIPKPDGLRPIVMDYVGA 642
633 RALGRKQAOHFTORLKTFLSMNRYERTKPHLMGSSVLGMNDIYRTWRAFLVRVALDQ 692
643 RTFREKRAEELTRVKALFSVLNFEARPGILGASVLGLDDIHRAWRTFVLVRQDP 702
693 TPRMYFVKADVTGAYDAIPOKGLVEVVMNRHSESTYCIQVAVVRDSQGOVHKFSFR 752
703 PPELYFVKVDVTGAYDTIPQDLRTEVIAIIPKNTYCVRYAVVQKAAHGHVRKAFKS 761
753 QVTTLSLQPMYMGFLKHLODSASALRNSVITEOSTSMNESSSLFDFLHFLHSVK 812
762 HVSTLTDLPQNRQFVAHQET--SPURDAVVEQSLSNEASGLFDVFLRFMCHAVR 819
813 LGDRCYTCQGIPOGSSLSLTLSCFCGDMENKLFABEVORDGLLRFVDDFLVTPHLDQ 872
820 IRGKSYVQCQGIPOGSSLSLTLSCYCGDMENKLFAGIRRDGLLRLVDDFLVTPHLDQ 879
873 AKTFLSTLVHGVPEYCGMINOKTVNFPVPEPTLGGAAPOYPAHCLFPWCGLLDITOT 932
880 AKTFLRTLVRGVPEYCGVNVNLRKTVNFPVPEDEALGATFVQMPAHGLFPWCGLLDITRT 939
933 LEVFCDSYGAQTSIKTSLTPQSVFKAGTKMRNKLKLSVLRLKCHGLFLDLQVNSLQTVCI 992
940 LEVQSDYSYARTSIRASLTNRGFKAGRNWRKLFGLVRLKCHSLFLDLQVNSLQTVCT 999
993 NIYKIFLOQVRFHACVQLQFPDORVRKNLTFGLIISQASCCYAILKVNPGWTLKAS 1052
1000 NIYKILLQVRFHACVQLQFPHQVQWKNPTFFLRVSDTASLCSYLKAKNAGMSLGAH 1059
1053 GS--FPPAARHLCYQAFLLKLAHSHVYIKLLGLPLRTAQKLLCKRLPEATWTILKAAA 1109
1060 GAAGPLPSEAVQWLCHQAFLLKLTFRHVTYVPLLGSLRTAQTLQSLRKLPGTTLTLEAAA 1119
1110 DPALSTDFQTILD 1122
1120 NPALPSDFKTILD 1132

```

SULT 8

-08-912-951-2  
Sequence 2, Application US/08912951  
Patent No. 6475789

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.

```

APPLICANT: Morin, Gregg B.
APPLICANT: Hanley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
TITLE OF INVENTION: THERAPEUTIC METHODS
NUMBER OF SEQUENCES: 335
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/912,951
FILING DATE: 14-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00260005
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1132 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-912-951-2

```

Query Match 59.4%; Score 3505; DB 4; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

QY 1 MTRAPRCPAVRSLLRSRYREVWPLATFVRRLQPEGRRLVQPGDPKXYRTLVACQLVCMHW 60
DB 1 MPRAPRCRAVRSLLRSRYREVLPATFVRRLQPGQWRLVQRGDPAAFRALVAQCLVQCVFW 60
QY 61 GSQPPPADLSFHOVSSIKELVARVVQRLCERNVRLAFGFELLNEARGGPPMAFTSSVR 120
DB 61 DARPPPAAPSFQVSCIKELVARVVQRLCERNVRLAFGFALLDARGGPPMAFTSSVR 120
QY 121 SYLPTVETLRVSGAWMLLSRVGDDLLVYLLAHLCALYLLVPPSPCAYQVCGSPLYQICA 180
DB 121 SYLPTVTDALRGSGAWGLLLRRVGGDVLVHLLARCAFLVLPVAPSCAYQVCGSPLYQLGA 180

```

```

181 TTDIWPSVGSASVPTBPVGRNFTNLFLOQIKSSRQEQAPKDLALPSEGTGKXHLSTSTS 240
181 ATQARPPPHAS-GFRRRLG-----CERAWNHSVREAGVPLGLPAPCARRRGCSASRS 231
241 VPSAKKARCVPVPRVEEGP-----HRQVLTPTGKSN-VSPARSPEVPTAEKDLSSK 292
232 LPLPKPRCAAEPEPTPVGGQSWAHPORTGPDGRGCVSPAR-----PARENTSLE 286
293 GKVSJDSLGS-GSVCCGKHSSTLSLSPQNAPQLRP-FIETRHPLYSRGQGRLNPSF 350
287 GALSGRHSHSPSVGRGHAGPSTSPRPMDTPCPVVAETKHPFLYSYG-D-KEQLRPSF 345
351 LLSNLOPNTGARELVEIIFLGRSPRTSPLCRTHLSRYWQMPLEFQQLLVNHAEOY 410
346 LLSLSPSLTGARELVETIFLSRPMWPTPRLPDORYWQMPLEFLLGNHGAQCPY 405
411 VLLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSPMWVY 452
406 GVLLKTHCPURAA---VTPAAGVCAREKPGSVAAPEEDTDPRLVOLLRQHSGSPWVY 462
453 GFRLACLVKTVSASLNGTRHNERFPKNLKKETISLGKYKLSLOELVMKMKVEDCHILRS 512
463 GFVACLRLRVPLGMSRHRNRRFLTKKTFISLGKHAQLSQELTWMKSVRDCAWLRR 522
513 SPGRDVPAAEHLRERILATLFLWMDTVVYVQLLRSFFYITESTFQKNRFFYKKSVM 572
523 SPGVGCVPAEHLRERILAKFLHMLMSVYVVELLRSFFYVTTTFQKNRFFYKKSVM 582
573 KLSIGVROHLRVRARELSOBRVHQDTWLMPICRLPRTPKNGRLRPVNMYSYGT 632
583 KLSIGVROHLRVRARELSOBRVHQDTWLMPICRLPRTPKNGRLRPVNMYSYGT 642
633 RALGRKQAOHQFTORLKTIFSMNLNYERTKPHLMGSSVLGMNDIVRTWRAFVLRVRLDQ 692
643 RTFEREKARELTSRVKALFSVNLNREARPGLLGASVLGLDDIHRAMRTFVLRVRAQDP 702
693 TPRMYFVADVTGAYDAIPQKLUVEVNVAMNIRHSSTYCIROYAVVRDSDQGVKHSFR 752
703 PPELYFVKVDVTGAYDTIPQDLRTVEIASIIL-PQNTYCVRRYAVVQKAAHGHVRKAFKS 761
753 QVTTLSPLQPMQGLKHLQSDASALRSVVIQSIWNESSSISLDFLFLHLSHVVK 812
762 HVSTLIDLPQYMQFVAHQET--SPERDAVIEQSSLINEASSGLFDVLFKCHHAVR 819
813 IGRCYTPCCGIPQGSLSLTLCSLCGDMENKLFAGVQDGLLRLFDVDDFLVTPHLDQ 872
820 IRGKSYVCCGIPQGSLSLTLCSLCGDMENKLFAGVQDGLLRLFDVDDFLVTPHLDQ 879
873 AKTFLSLVHGVPEYGCMLNLTQVNVFPVPGTIGGAAPYQLPAHCLPPWGLLLDTQT 932
880 AKTFLSLVHGVPEYGCMLNLTQVNVFPVPGTIGGAAPYQLPAHCLPPWGLLLDTQT 939
933 LEVFCDSYGAQTSIKTSTFQGVFKAGTMKNKLSVLRKCHGLFLDLQVNSLQTVCI 992
940 LEVQSDYSSVARTSIRASLTFRNGFKAGRMKRLFGVLRKCHSLFLDLQVNSLQTVCI 999
993 NIYKIFLQAYRHACVQLPFPQVRKNTLFPGLIISQAQSCCVAILKVNPGMTLKAAS 1052
1000 NIYKIFLQAYRHACVQLPFPQVRKNTLFPGLIISQAQSCCVAILKVNPGMTLKAAS 1059
1053 GS---FPPEAAHMLCYOAPLKLAAHSLVYKCLGLPLRTAQLKCRKLPEATMTILKAAA 1109
1060 GAAGPLPSEAVQMLCHQAFLLKLTTRHVTYVPLGLSLRFTAQLSKLPFGTTLTAAEAAA 1119
1110 DPALSTDFQITLD 1122
1120 NPALPSDFKITLD 1132

```

ULT 9  
08-974-549A-611  
sequence 611, Application US/08974549A  
atent No. 6166178

```

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin B.
APPLICANT: Andrews, William H.
TITLE OF INVENTION: Human Telomerase Catalytic Subunit
NUMBER OF SEQUENCES: 727
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974.549A
FILING DATE: 19-NOV-1997
CLASSIFICATION: 536
APPLICATION DATA:
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/851,843
FILING DATE: 06-MAY-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/854,050
FILING DATE: 09-MAY-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/911,312
FILING DATE: 14-AUG-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/912,951
FILING DATE: 14-AUG-1997
APPLICATION DATA:
APPLICATION NUMBER: US 08/915,503
FILING DATE: 14-AUG-1997
APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17618
FILING DATE: 01-OCT-1997
APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/17885
FILING DATE: 01-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph Ted
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002610US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 611:
SEQUENCE CHARACTERISTICS:
LENGTH: 1154 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1154

```



OTHER INFORMATION: /note= "fusion protein composed of hTERT  
OTHER INFORMATION: protein sequence, vector sequences, the  
OTHER INFORMATION: Myc epitope and His6 tag"

08-974-549A-611

Query Match 59.4%; Score 3505; DB 3; Length 1154;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
1 MTRAPCPAVASLLRSRYREYVPLATFVRRLGPEGRELVQGPDKIYRILVAOCLVCMFW 60  
1 MPRAPCRVAVSLLRSRYREYVPLATFVRRLGPEGRELVQGPDKIYRILVAOCLVCMFW 60  
61 GSQPPADLSHOWSSILKELVAVVQRLCERNERNVLAFLNENARGGPPNAFTSSVR 120  
61 DARPPAPAPSPFRQVSCLELVARVLQRCERGAKNVLAFLNENARGGPPNAFTSSVR 120  
121 SYLPTNVTIETLRVSGAMMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLPTNVTIETLRVSGAMMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASVYRTPVGRNFTNLRFLQIKSSRQBAKPLALPSRGTKYRHLSTSTS 240  
181 ATQARPPPHAS-GERRRLLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
241 VPSAKKARCPVPRVEEGP-----HRQVLTPTSGKSW-VPSPARSPVPTAEKDLSSK 292  
232 LPLPKRGAAPERPFPVQGSWAHGRTPGSDRGFCVSPAR-----PAEATSLE 286  
293 GKVDLSLS-GSVCKKHPSTSLSPRQNAFQIRP-FIETRHFLYSRGDGERLNPSF 350  
287 GALSGRHSHPVGRHQHAGPPTSRPRPMDTFCPPVYAEKTHFLYSSGD-KEQLRPSF 345  
351 LLSNLPQNLTGARRLVEIFLGRSPTSGPCRTHLSRRYQWQRPFLQOLLVNHAEQV 410  
346 LLSLRPSLTGARRLVEIFLGRSPTSGPCRTHLSRRYQWQRPFLQOLLVNHAEQV 405  
411 VLLRSHCRFRFANQCVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
406 GVLLKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEEDTPRRLVQLLRQHSWPQVY 462  
453 GLRLACLVLGASLWGRHNRERFKNKFIKGLYKGLSLOELAKMKVDCDCHWLS 512  
463 GVRACRLVLPVPGWGRHNRERFKNKFIKGLYKGLSLOELAKMKVDCDCHWLS 522  
513 SPGRKVPAAEHLRERILATFLWIMTYVQLRSFFYITESTFQKRLFFYRKSWS 572  
523 SPGVGCVPAEHLRERILATFLWIMTYVQLRSFFYITESTFQKRLFFYRKSWS 582  
573 KLSIGVROHLERVLRLSQQEVRHODTWLAPICLRLFPKPNGLRIVNMSYSGT 632  
583 KLSIGVROHLERVLRLSQQEVRHODTWLAPICLRLFPKPNGLRIVNMSYSGT 642  
633 RALGRKQAOHTQRLKTLFSLMNTYRTKPHLMGSSVLGNMDIYRTWRAFVLRVLDQ 692  
643 RIFRREKRAERLTSRVKALFVLNVERAERFGLLGASVLGLDDHRAWRFTVLRVADP 702  
693 TPMPYFKADVTDGAYDAIQGKLVNVMNIRHSESTYCIQYAVWRDQSGQVHKFR 752  
703 PPPLYFKVDVTGAYDTIPQDLRLTVEIASIK-PQNTYCVARYAVQKAAGHVRKAFS 761  
753 QVTTLSDLQPYMGQFKHQDSDASALRNSVVEIOSISMNESSSLDFFLHFRHSVVK 812  
762 HVSTLTDLPYMQFVAHQET--SPLRDVAVIEQSSSLEASSGLFDVFLRFCHAVR 819  
813 IGRCYTCQGIPOGSSSLTLLCSLCFGDMENKLPFAVQVRDGLLLRFVDDFLVTPHLDQ 872  
820 IRGKSVYVQGIPOGSSSLTLLCSLCFGDMENKLPFAVQVRDGLLLRFVDDFLVTPHLDQ 879  
873 AKTFELSLVHGVPEYCCMNLQKTVNVEPFTLGGAPYQLPAHCLFPWCGLLDITQT 932  
880 AKTFELSLVHGVPEYCCMNLQKTVNVEPFTLGGAPYQLPAHCLFPWCGLLDITQT 939

QY 933 LEVPCDYSYGAQTSIKTSLTFQSVFKAGTKVKNKLLSVLRKCHGLFLDLQVNSIQTVCI 992  
DB 940 LEVQSDYSYGAQTSIKTSLTFQSVFKAGTKVKNKLLSVLRKCHGLFLDLQVNSIQTVCI 999  
QY 993 NIYKIFELQAYRPHACVIOQLPFDQVRKOLTFEFLGIISQASCCYAILKVKXNPGMTLKAS 1052  
DB 1000 NIYKIFELQAYRPHACVIOQLPFDQVRKOLTFEFLGIISQASCCYAILKVKXNPGMTLKAS 1059  
QY 1053 GS---PPPEAAHWCQAFELKLAHSHVYIKCLGLPLTAQKLCGLPEATMTTLKAAA 1109  
DB 1060 GAAGPUPSEAVQWMLCHQAFLLKLTNRVTVYVLLGSLRTAQQLSRKLPFGTTLTAAEAAA 1119  
QY 1110 DPALSTDFOTILD 1122  
DB 1120 NPALPSDFKILD 1132  
RESULT 10  
US-08-912-951-323  
Sequence 323, Application US/08912951  
Patent No. 6475789  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Hatley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
TITLE OF INVENTION: THERAPEUTIC METHODS  
NUMBER OF SEQUENCES: 335  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/912,951  
FILING DATE: 14-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002600US  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 323:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

-08-912-951-323

Query Match 59.4%; Score 3505; DB 4; Length 1154;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVRSILSRVREVPVPLATFVRRLPGSGRRLLVQDPKIVRTLVLAQCLVCMHW 60  
1 MTRAPCPAVRSILSRVREVPVPLATFVRRLPGSGRRLLVQDPKIVRTLVLAQCLVCMHW 60  
61 GSOPPPADLSFHQVSSKLVARVVRQLCERNRNVLAFGPELLNEARGGPPMAFTSSVR 120  
61 DARPPAASFPQVCLKELVARVLQRCERCAKNVLAFGELLDDGARGGPPMAFTSSVR 120  
121 SYLNPVTITLVSGAMMLLSRGDDLLVLLAHALYLLVPPSCAYQVCGSPLOYICA 180  
121 SYLNPVTITDALRGSGAWGLLRVGDVLLVHLLARCALFVLVAPSCAYQVCGSPLOYICA 180  
181 TTIDVPSVASYRPTPVGRNFTNLRFQIKSSSQEAPKPLALPSRGTKRHLSTSTS 240  
181 ATCAQPPPHAS-GPERRIG-----CERANHSVRGAGVPLGLPAPGARRGGASRS 231  
241 VPSAKKACRCYPRVVEEG-----HRQVLTSPGKSW-VPSPARSPVPTAKDLSSK 292  
232 LELPKRPRGAAPERTPVQGSWAHPGRTGSDRGFCVVSAPAR-----PAEAEATSL 286  
293 GKVSDDLSS-GSVCKHKPSSTLSLSPQNAFQLRP-FIETRHLYSRGDDGOERLNPSP 350  
287 GALSGRHSHPSVGRQHAGPPSTSRPRPMDTPCPVYAKTHFLYSSGD-KEQURPSF 345  
351 LLSNLPQNLGARRIVEIFLGSRRPTSGPLCRTHLSRRYQWRPLFQQLVNHABCOY 410  
346 LLSLRPLSITGARRLVEITFLGSRPMPGTPRRLPRLPQRYQWRPLFLELGNHACPY 405  
411 VLLRSHCHFRFANQOVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
406 GYLLKTHCPLRAA---VTPAAGVCAKREPOGSVAAPEDDTPRRLVQLLRQHSPPWQV 462  
453 GFLRACLCKVSGASLWGRTHNRERFFKNLKFISLGKYKLSLOELMWKMKVEDCHWLS 512  
463 GFVRACLRRLVPPGLWGRHNRERFLNTKFIISLGRHAKLSLOELTWKMSVRDCAWLR 522  
513 SPGRORVPAEHRRLERILATFLFWMPTVYVQLLRSFFYITESTFQKRLFPYKSVWS 572  
523 SPGVGCVPAEHRRLERILAKFLHLMVSVYVVELLSFFYVTTFTFQKRLFPYKSVWS 582  
573 KQSGVGRHLRRLRLSQEVEVHHQDTWLAMPICRLRIPKNGRLRIVNMSYSGT 632  
583 KQSGIGIRHLKRVQLRSLSAEVQHQREAPALLTSRLRIPKDPGLRIVNMDYVYGA 642  
633 RALGRKRAQHFTORLKTFLSMYNTERTKPHLMGSSVLGMNDIYRTWRAFLVLRALDQ 692  
643 RTFRKRAERLTSVKALFSLVNYERARRGGLLGASVLGLDDIHRARWTFVLVRAQDP 702  
693 TRMVFYKADVTGADATPOGKLVEVYVNMTHSESYCIROYAVVRDQSGQVHKSFRR 752  
703 PPFLYFVKVDVTGADTIPQDLRTVIAIISK-PQNTYCYVRYAVVQKAAGHVRKAFKS 761  
753 QVTTLSDLQPYMGQFLKLODSDALNSVSVIEQSIEMNESSSLDFFLHLRHSVVK 812  
762 HVSITLTLQPYNQVFAHQET--SPLDADVIEQSSSLNEASSGLFDVFLRMCWHAVR 819  
813 IGRCYTCQCGIPQSSSLTLLCSLCFGDMENKUPAEVQRDGLLRLFRVDDFLVTPHLDQ 872

## RESULT 11

US-08-974-549A-613

; Sequence 613, Application US/08974549A

; Patent No. 6166178

; GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.

; APPLICANT: Lingner, Joachim

; APPLICANT: Nakamura, Toru

; APPLICANT: Chapman, Karen B.

; APPLICANT: Morin, Gregg B.

; APPLICANT: Harley, Calvin B.

; APPLICANT: Andrews, William H.

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit

; NUMBER OF SEQUENCES: 727

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA: US/08/974,549A

; FILING DATE: 19-NOV-1997

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/724,643

; FILING DATE: 01-OCT-1996

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/844,419

; FILING DATE: 18-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/846,017

; FILING DATE: 25-APR-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/851,843

; FILING DATE: 06-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/854,050

; FILING DATE: 09-MAY-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/911,312

; FILING DATE: 14-AUG-1997

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 613:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1189 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1189  
OTHER INFORMATION: /note= "fusion protein composed of  
OTHER INFORMATION: htrt signal sequence and full length  
OTHER INFORMATION: htrt protein"  
3-08-974-549A-613

Query Match 59.4%; Score 3505; DB 3; Length 1189;  
Best Local Similarity 62.4%; Pred. No. 0;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
/ 1 MTRAPCAVRLSLRLRYREVMPLATVRLRGPEGRRLVQDPKLYRTLVACQVLCVHMW 60  
/ 58 MPRAPCAVRLSLRLRYREVLPLATVRLRGPEGRRLVQDPKLYRTLVACQVLCVHMW 117  
/ 61 GSQPPADLSFQVSSKELVARVQVQLCERNRNVLAFGLFELLNFAARGGPMATSSVR 120  
/ 118 DARPPAPSPRQVSCVSKELVARVQVQLCERNRNVLAFGLFELLNFAARGGPMATSSVR 177  
/ 121 SYLNTVITLVRSGAWMLLSRVGDDLLVYLLAHCALVLLPSPCAVQVCGSPLYQICA 180  
/ 178 SYLNTVITLVRSGAWMLLSRVGDDLLVYLLAHCALVLLPSPCAVQVCGSPLYQICA 237  
/ 181 TTDWPSVSAVRYTRPVGRNFTNRLQKSSRCQKAPKPLALPSRGTKXHLSTSTS 240  
/ 238 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRGSGSRS 288  
/ 241 VPSAKKARCYVPVRVERGP-----HRQVLPSPGKSM-VPSPARSPVPTAEKDLSSK 292  
/ 289 LPLKCPRRGAAPERTPVQGSWAHPGTRGSDRGFCVVSPPAR-----PAEATSLE 343  
/ 293 GKVDLSLS-GSVCKKHPSPSTLLSPRONAQALRP-FIETHFLYRGDQGRINPSF 350  
/ 344 GALSGRTHSPSVGRQHAGPPSTRPPEPMDTPCPPEVAETKHFYSSGD-KEQLRPSF 402  
/ 351 LLSNLQPNLTGARLVELIFLGSPPRTSGPLCRTHLSRRYQWMPFLFOOLLVNAECQY 410  
/ 403 LLSNLQPNLTGARLVELIFLGSPPRTSGPLCRTHLSRRYQWMPFLFOOLLVNAECQY 462  
/ 411 VRLLRSHCRFTANQOVTDAL-----NTSPHMLDLRLHSSPWQVY 452  
/ 463 GVLLKTHCPRAA---VTPAAGVCAKPKQGSVAAPPEEDTPRRLVQLLRHSSPWQVY 519  
/ 453 GFLRACLCVYSASLWGTNRHNERFFNKKFSLGKYKLSLQELMWKXVEDCHLRS 512  
/ 520 GFVRACLRLVPPGLWGRSHNRERFLNKKFSLGKYKLSLQELMWKXVEDCHLRS 579  
/ 513 SPGKDRVPAEAHRLRERLATFLFWMMDTVVYVQLLRSPFFYITESTTQKNRFLFYRKSVMS 572

DB 580 SPGVGCVAAAEHRLREELAKPLHLMVSVYVVELLRSFFVTETTFQKNRFLFYRKSVMS 639  
QY 573 KLSQIGVROHLERVELRELSEVEVRHQDTWLMPIICRLRFIPKPNGLRPIVNNYSYSGMT 632  
DB 640 KLSQIGIRGHLKRVQLRELSEAEVRQREARPAALLTSRLAFIPKPDGLRPIVNDYVYVGA 699  
QY 633 RALGRKQAQCHFTORLKTLSMAYVERTKPHLMGSSVLGNNDIYRTWRAPVLRVLRALDQ 692  
DB 700 RTFRREKRAERLTSRVKALFSLVNYERARREGGLGASVIGLDDIHRARFTFLVRRAQDP 759  
QY 693 TPRMYFVKADVTGAYDAIPQGLVEVWAMIRHSESTYCIQYAVVRRDSQGVYHKSFR 752  
DB 760 PRELYFVKVDTGAYDTIPQDLRETVASIIK-PQNTYCVRYAVVQKAAHGHVRKAFKS 818  
QY 753 QVTTLSDIOPYMGQFLKHLQSDASALRNSVVIQSSIMNRESSSLDFLHFLRHSVVK 812  
DB 819 HVSTLTDLQVYRQFVAHQET--SPLRDAVVIQSSSLNEASSGLFDVFLRFMCHHAVR 876  
QY 813 IGRDVCYTCQGIPOGSSILSTLLCSLCFGDMENKLFQVQDGLLRFVDDFLVTPHLDQ 872  
DB 877 IRGKSVYQCGIPQSSILSTLLCSLCYGMENKLFAGIRRDGLLLRLVDDFLVTPHLDQ 936  
QY 873 AKTFLTLVHGVPEYGCINLQKTVVNFPEPGLTGAAPYQLPAHCLFPWCGLLLDTOT 932  
DB 937 AKTFLTLVHGVPEYGCINLQKTVVNFPEPGLTGAAPYQLPAHCLFPWCGLLLDTOT 996  
QY 933 LEVFCDSYGAOTSITKSLTQSVFKAGTKWRNKLTVLRKCHGLFLLQVNSLQVTCI 992  
DB 997 LEVQSDSYSTARSITRNLNFGKAGNRNKLFGVLRKCHGLFLLQVNSLQVTCI 1056  
QY 993 NIYKIFLQAYRHACVILQLPDQVRNKLTFGLITSSQASCCYAILKVNQPMWLKAS 1052  
DB 1057 NIYKIFLQAYRHACVILQLPDQVRNKLTFGLITSSQASCCYAILKVNQPMWLKAS 1116  
QY 1053 GS---FPPAAHWLCVQAFLLKLAHSLVYKCLGLPRTACKLCKRLPEATMTILKAAA 1109  
DB 1117 GAAGPLPSEAVOMLCHQAPFLKLRHRYTVYVPLGLSRLTAQQLSRKLPGLTILALEAAA 1176  
QY 1110 DPALSTDFQITILD 1122  
DB 1177 NPALPSDFKTLID 1189  
RESULT 12  
US-08-912-951-325  
; Sequence 325, Application US/08912951  
; Patent No. 6475789  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Liringer, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND  
; NUMBER OF SEQUENCES: 335  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/912,951

FILING DATE: 14-AUG-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/851,843  
 FILING DATE: 06-MAY-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002800US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 325:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1189 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 08-912-951-325

Query Match 59.4%; Score 3505; DB 4; Length 1189;  
 est Local Similarity 62.4%; Pred.No. 0;  
 matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCPAVSLRLSRVYRWPLATVRLGEGRLVQGPDKIYRLLVAQCLVCHW 60  
 58 MPRAPCPRAVSLRLSRVYRWPLATVRLGEGRLVQGPDKIYRLLVAQCLVCHW 117  
 61 GSQPPADLSFHYSSLLKELVARVVOQLCERNERNVLAQFELLNEARGGPPMAFTSSVR 120  
 118 DARPPAPSPFQVSCILKELVARVVOQLCERNERNVLAQFELLNEARGGPPMAFTSSVR 177  
 121 SYLNTVITLRLVSGAMWLLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPYQICA 180  
 178 SYLNTVITLRLVSGAMWLLLSRVGDDLLVLLAHALYLLVPPSCAYQVCGSPYQICA 237  
 181 TTDIWPVSASVYRTPVGRNFTNLRFLOQIKSSRCEAPKPLALPSRGTKRHLSTSTS 240  
 238 ATQAPPPHPS-GRPRUG-----CERAWNSVREAGVPLGLFAPGARRGGASRS 288  
 241 VPSAKKACVYVPRVEEGP-----HRQVLPFGKSM-VPSAPRSPEVPTAEKDLSSK 292  
 289 LPLKPRRGAAPPEPTEPVQGGSWAHFGRTRGSDRGFCVWSFAR-----PAEAEATSL 343  
 293 GKVSIDLSS-CSVCKKHPSTSLSPRQNAQLRP-FLETBFLYLSRGDGOERLNPSP 350  
 344 GALSCTRHSFVSQRHAGPPSTSPRPRPDWTPCPVYAEETKHFLYSSGD-KQELRPSF 402  
 351 LLSNLQPNLTGARRLVEIFLFGSPRTSGPLCRTHLSRRYQWRPLFQQLLVNHAECY 410  
 403 LLSLSRPSLTGARRLVEIFLFGSPRTSGPLCRTHLSRRYQWRPLFQQLLVNHAECY 462  
 411 VLLRSHCRFTANQVTDAL-----NTSPHMLMDLLRHSSPWQY 452  
 463 GVLLKTHCFLRAA---VTPAAGVCAKPGQSVAAPEEDTDPRLVQLLRQSSPWQY 519

QY 453 GFLRACLCKVVSASLWGTNRHNERFFKNLKKFISLGGYKGLSLOELAMWKKVEDCHWLS 512  
 DB 520 GFVRACLRRLVPPGLWGRHNERRLNRTKKFISLGGHAKLSLOELTWKMSVRDCALWR 579  
 QY 513 SPQKRVPAAEHRLRERILATFLWLMNTVYVQLRSFFYITESTFQKNLFFYRKSVWS 572  
 DB 580 SPQVGVCPAAEHLREEILAKFLHLMVSVYVVELLRSPFYVTTFTFQKNLFFYRKSVWS 639  
 QY 573 KLOSIGVROHLERVLRELQBEVRRHODTWLAPICELRIKPKNGRLRPVNMYSMTG 632  
 DB 640 KLOSIGIRQLKRVQLRELSEABVQRHREARPAULTSLRIFPKPDGLRPVNMVYVGA 699  
 QY 633 RALGRKQOACHETQRLKTLFSLMYNERTKHPHMGSSVLGMDIYRTWRFAVLVRALDQ 692  
 DB 700 RTRPREKRAERLTSRVKALFSLVNYERARRPGLLGASVLGLDDIHRAWRTVLVRAQDP 759  
 QY 693 TPRMYFVKADVTGAYDAIPOGKLVEVVMNMRHSESTYCIROYAVWRDSDGQVHKSPFR 752  
 DB 760 PPELYFVKVDVTGAYDTIPQDLTEVIASIIK-PQNTYCVRRYAVVQKAAHGHVYRKAFKS 818  
 QY 753 QVTTLSLOPYMGQFLKHLQDSASALRNSVVIEOSISMNESSSLDFDFLHFLRHSVVK 812  
 DB 819 HVSTLTLQFYRQFVAHLQET--SPLRDVVIEOSSSLNEASSGLDFVFLRFCHAVR 876  
 QY 813 IGRCYTCQCGIPQGSSTLTLCSLCFGDMENKLPFAVQORDGLLLRFVDDFLVTPHLDQ 872  
 DB 877 IRGKSVVQCGIPQGSSTLTLCSLCYGDGMENKLPFAGIRRDGLLLRLVDDFLVTPHLDQ 936  
 QY 873 AKTFLSTLVHGVPEYCGMINLOKTVNRPVPEBPTLGGAPYQLPAHCLFPMWGLLDTQT 932  
 DB 937 AKTFLTLVRGVPEYCGVNLKTVNFPVEDEALGGTAFVQMPAHGLFPMWGLLDTQT 996  
 QY 933 LEVFCDSYGAOTSIKTSLTFQSVFKAGTKMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 DB 997 LEVQSDYSYVARTSIRASLTFNRGFKAGNWRKLFVLRKCHSLFLDLQVNSLQTVCT 1056  
 QY 993 NIYKIFLOAYRPHACVIOLPDQVRKNLIFPLGIISSQASCCYAILKVKNGPMTLXAS 1052  
 DB 1057 NIYKILLQAYRPHACVIOLPDQVRKNLIFPLGIISSQASCCYAILKVKNGPMTLXAS 1116  
 QY 1053 GS---PPPEAAHWLCYQAFLLKLAHSAVYKCLLQPLRTAQKLCRKLPEATMTILKAAA 1109  
 DB 1117 GAAGPLPSEAVQWLCQAFLLKTHRTVTVYVPLLSLRTAQKLCRKLPEATMTILKAAA 1176  
 QY 1110 DPALSTDFOTILD 1122  
 DB 1177 NPALPSDFKTILD 1189

## RESULT 13

US-08-974-549A-612  
 ; Sequence 612, Application US/08974549A  
 ; Patent No. 6166178

## GENERAL INFORMATION:

; APPLICANT: Cech, Thomas R.  
 ; APPLICANT: Linsner, Joachim  
 ; APPLICANT: Nakamura, Toru  
 ; APPLICANT: Chapman, Karen B.  
 ; APPLICANT: Morin, Gregg B.  
 ; APPLICANT: Hatley, Calvin B.  
 ; APPLICANT: Andrews, William H.  
 ; TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
 ; NUMBER OF SEQUENCES: 727  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Townsend and Crew LLP  
 ; STREET: Two Embarcadero Center, Eighth Floor  
 ; CITY: San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM: disk  
 ; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17685  
FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 612:

SEQUENCE CHARACTERISTICS:

LENGTH: 1200 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Protein

LOCATION: 1..1200

OTHER INFORMATION:

/note= "fusion protein composed of His6

and Anti-Xpress tags, enterokinase

OTHER INFORMATION: cleavage site and full length hprt

OTHER INFORMATION: protein

;08-974-549A-612

Query Match 59.4%; Score 3505; DB 3; Length 1200;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPRCVRSLLRSYREVWPLATFVRLGEGRLVQDPDKIYKTLVAQCILVCMHW 60

69 MPRAPRCVRSLLRSYREVWPLATFVRLGEGRLVQDPDKIYKTLVAQCILVCMHW 128

61 GSOPPPADLSFHQVSSIKELVARVQRLCERNERNVLAFCFELLNEARGCPMAFTSSVR 120

129 DARPPPAPEFRQVSCIKELVARVQRLCERNERNVLAFCFELLNEARGCPMAFTSSVR 188

121 SYLNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALVLLVPPSCAYQVCGSPLYQICA 180

Db 189 SYLNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALVLLVPPSCAYQVCGSPLYQICA 248  
QY 181 TTDIWPVSASYPTRFVGRNFTNLRFLOQIKSSSSOEAPKFLALPSPRTKXHLSTSTS 240  
Db 249 ATQARPPPHAS-GPRRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGSASRS 299  
QY 241 VPSAKKARCYPVRVEEGP-----HRQVLTPSGKSW-VSPASPEVPTAEKDLSSK 292  
Db 300 LPLPKRPRRGAAPPEFPTVGOQSWAHGRTGTRGSDRGFCVVSFAR-----PABEATSL 354  
QY 293 KVSVDLSLS-GSVCKHKPSSTLSLSPRONAFQLRP-FIETRHFLYSRGDQGERLNPSF 350  
Db 355 GALSOTRHSHPVGRQHGAGPSTSRPPRPWDTPCPVVAETKHFLYSSGD-KEQLRPSF 413  
QY 351 LLSNLQPNLIGARLVEIILFSGRPRTSGPLCTHLSRRYQWOMRPLFOQLLVNHAECY 410  
Db 414 LLSLRPSLTGARELVTIFLGRPMWPGTPRRLPRLPQRYWQMRPLFLELGNHQAQCY 473  
QY 411 VRLLRSHCRPTANQOVTDAL-----NTSPPHLMDLLRLHSSPWQYV 452  
Db 474 GVLLKTHOPLRAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRRLVQLLRQHS 530  
QY 453 GFLRACLCKVVSASLNGTRNERRFFXNKKFISLKYKGLSLOELMWKQVEDCHWLRS 512  
Db 531 GFVRACLRLVPPGLGSRNERRRFLNFKFISLGHAKLSLOELTWKSVRDCAWLR 590  
QY 513 SPGRDVPAAEHRRLERILATFLFWMMDTVVOLLRSFFYITESTFQKRLFFYRSVMS 572  
Db 591 SFGVGCVPAAEHRRLERILAKFLHLSMVVVELLSFFVYVTTTQKRLFFYRSVMS 650  
QY 573 KLOSTGVROHLERVLRELSEGEVRRHODTWLAMPICRLRPIPKPKNGLRPIVMSMGT 632  
Db 651 KLOSTGIRQHLKRVQLRELSEAEVRRHODTWLAMPICRLRPIPKPKNGLRPIVMSMGT 710  
QY 633 RALGRKQAOHFTORLKTLSMLNVERTKHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
Db 711 RTFREKKAERLTSRVKALFSLNTERARPPGLGASVLGLDDIHRARWTFVLVRADP 770  
QY 693 TPRMYFVKADVTGAYDAIPQKLVVAVNMRHSESTYCIQAVVVRDQSGQVHKSFR 752  
Db 771 PPELYFVKVDVTGAYDITPDRLTEVIAIIPQNTYCVRYAVVQKAAHGHVRAKFS 829  
QY 753 QVTTLSLQPYMGOLFHLQSDASALRNSVLIQSIQSMNNESSSLDFDFLHFLRHSVVK 812  
Db 830 HVSITLTDQPYMRQFVAHLQET--SPLRADAVVIEQSSSNEASSGLFDVLRFAHCHAVR 887  
QY 813 IGDRCYTCQGIPOGSSLSLTLCSLCFQDMENKLFQVQVQDGLLLRPFVDDFLVTPHLDQ 872  
Db 888 IRGKSYVQCGIPOGSSLSLTLCSLCYQDMENKLFQVQVQDGLLLRPFVDDFLVTPHLDQ 947  
QY 873 AKTFLSTLVHGVPEYGCMINLQKTVNPFVPEPTLGGAAVQLPALHCLFPWCGLLDTOT 932  
Db 948 AKTFLSTLVHGVPEYGCVINLQKTVNPFVPEPTLGGAAVQLPALHCLFPWCGLLDTOT 1007  
QY 933 LEVFCDSYGAQTSIKTSITFQSVFKAGKTMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
Db 1008 LEVQSDYSYARTSIRASITFNRGFKAGNRRKLFGLVLRKCHSLFLDLQVNSLQTVCI 1067  
QY 993 NIYKFLQAYRFHACVQLQDPQVRKXNLTFFLGIITSSQASCCYAILKVNPGMTLKAS 1052  
Db 1068 NIYKILLQAYRFHACVQLQDPQVRKXNLTFFLGIITSSQASCCYAILKVNPGMTLKAS 1127  
QY 1053 GS---PPPEAAHLCVQAFLLKLAHNSVIYKCLLQPLRTAQKLCRKLPEATMTILKAAA 1109  
Db 1128 GAAGPLPSEAVNWLCHQAFLLKLTREHVTYVPLIGSLFETACTQLSRKLPGTTLTALEAAA 1187  
QY 1110 DPALSTDFQITLD 1122  
Db 1188 NPALPSDFXTILD 1200

RESULT 14

08-912-951-324

Sequence 324, Application US/08912951

Patent No. 6475789

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.  
 APPLICANT: Lingner, Joachim  
 APPLICANT: Nakamura, Toru  
 APPLICANT: Chapman, Karen B.  
 APPLICANT: Morin, Gregg B.  
 APPLICANT: Harley, Calvin  
 APPLICANT: Andrews, William H.  
 TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS  
 NUMBER OF SEQUENCES: 335  
 CORRESPONDENCE ADDRESS:  
 ADDRESSES: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/912,951  
 FILING DATE: 14-AUG-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/854,050  
 FILING DATE: 09-MAY-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002600US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 324:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1200 amino acids  
 TYPE: aming acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 08-912-951-324

Query Match 59.4%; Score 3505; DB 4; Length 1200;  
 Best Local Similarity 62.4%; Pred No. 0;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 1 MTRAPRCRAVRSLLRSRYREVMPLATFVRLGPGRRLVQPGDPKIVRTTLVAQCLVCMHW 60  
 69 MPRAFCRAVRSLLRSRYREVMPLATFVRLGPGRRLVQPGDPKIVRTTLVAQCLVCMHW 128

QY	61	GSOPPPADLSFHQVSSIKELVARVWORLCERNERNVLAFFGELLNEARGGPPMATTSVR	120
DB	129	DARPPAAPSPFRQVSCIKELVARVWORLCERNERNVLAFFGELLNEARGGPPMATTSVR	188
QY	121	SYLPTNVIETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPPSPCAYOVCGSPYQICA	180
DB	189	SYLPTNVIETLRVSGAWMLLSRVGDDLLVLLAHALYLLVPPSPCAYOVCGSPYQICA	248
QY	181	TTDIWPSVASYRPTRPVGRNFTNLRLQOIKSSSQEAPKPLAPSRGTKHLSTSTS	240
DB	249	ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGSASRS	299
QY	241	VPSAKKACYPVPRVEGP-----HRQVLPSPCKSW-VSPARSPEVPTAEKDLSSK	292
DB	300	LFLPRPRRGAAPERTPTVGQSWAHQRTGRGFCVVSAR-----PABEATSLE	354
QY	293	GKVDLSLS-GSVCCCKHKPSTSLSPPRQNAFLRP-FIETRHFLYSRGDQGERINPSP	350
DB	355	GALSGTRHSHPSVGRQHAGPPSTSRPPRWDTPCPVVAETKHFLYSSGD-KEQLRPSF	413
QY	351	LLSNLQNLTCARLVEIIFLGSRPRTSGPLCETHLSRRYQMRLFOOLLVNHAEQY	410
DB	414	LLSLRPSLTGARRLVETIFLGSRPMPGTPRLPELPORYQWMPFLFLELIGNHAQCPY	473
QY	411	VLLRSHCRERTANQOQVTDAL-----NTGPPHMDLLRLHSSPWQY	452
DB	474	GVLLKTHCPLEAA---VTPAAGVCAREKPGQSVAAPEBEDTDPRLVQLLRQHSPPWY	530
QY	453	GFLRACLCKVVSASLWGTNRNRRFPKNUKFTISLQYKGLSLQELMNMKVDECHWLS	512
DB	531	GFVRACLRLLVPPGLMGRSHNRERFLNTKFTISLGHAKLSLQELTWKMSVRDCAWLR	590
QY	513	SPGKDRVPAAEHRLREILATFLFWMMDTVVQLLSFFVITESTFQKNRLFYKSVMS	572
DB	591	SPGVGCVPAAEHRLREILAKFLHMLMSVYVELLSFFVITETIFQKNRLFYKSVMS	650
QY	573	KLQSIGVROHLRVLRELQSEVRHHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMT	632
DB	651	KLQSIGIRQHLKRVQLRELSEAEVROHREARPAALLTSRLRFIPKPDGLRPIVNM	710
QY	633	RALGRRKQAOHFTORLKLTFSLMNYERTKHPHLMGSSVLGMDIYETWEAFVLVRALDQ	692
DB	711	RTFRREKRAERLTSRVKALFSLVNIYERARRPGLLGASVLGLDDIHRAWTFLVLR	770
QY	693	TPRMVFKADVGTAYDAIPQGLVEVVVANNIRHSESTYCIROVAVVVRDSQGVHKSFR	752
DB	771	PPELYFVKADVGTAYDTIPQDLTEVIAIILK-PQNTYCVRAYVAVVQKAAHGVKAPKS	829
QY	753	QVTLSDLPYMGQFLKHLQDSASALRNSVITQISINWESSSLFDFFLFLRHSVVK	812
DB	830	HVSTLTLQPYMRQFVAHLOET--SPLRDVAVIEQSSSLNEASSGLFDVFLRPMCHAVR	887
QY	813	IGDRCYTCOCIGIPQSSSLTLLCSLCFGDMENKLPFAEVDORGLLRFDVDFLLVPHLDQ	872
DB	888	IRGKSYVQCGIPQSSSLTLLCSLCYGMENKLFAGIRRDGLLRLVDDFLVTFHLTH	947
QY	873	AKTFLSTLVHGVPEYGCMLNKQTVNVPVPEPTLGGAAAPYQLPAHCLFPWCGLLDTOT	932
DB	948	AKTFLRTLVRGVPYGCVNLKTVNVPVEDEALGTAFAVQMPAHGLFPWCGLLDTOT	1007
QY	933	LEVFDYSGYACTSIKTLTFQSVFKAGTKMRNKLKSLVRLKCHGLFLDQVNSLQTVCI	992
DB	1008	LEVOSDYSSVARTSIRASLTFNRGFKAGRNMRKLFGLVRLKCHSLFLDQVNSLQTVCT	1067
QY	993	NIYKIFLLQAVRFHACVIQLPFQQRVRKNLTFFLGIISSQASCCYAILKVRNPGMTLKAS	1052
DB	1068	NIYKIFLLQAVRFHACVIQLPFQQRVRKNLTFFLGIISSQASCCYAILKVRNPGMTLKAS	1127
QY	1053	GS----FPEPAHMLYLCQAFLKLAASHVYKLLGLPRTAQLCKLCKLPEATWTILKAAA	1109
DB	1128	GAAGFLPSEAVQWLCHOAFLKLTTRHRTVTVPLGLSLRTAQLCKLCKLPEATWTILKAAA	1187

1110 DPALSTDFQILD 1122  
: : : : :  
1188 NPALPSDFKILD 1200

## RESULT 15

08-974-549A-600  
Sequence 600, Application US/08974549A  
Patent No. 6166178

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.

TITLE OF INVENTION: Human Telomerase Catalytic Subunit

NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A

FILING DATE: 19-NOV-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph Ted

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 600:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1285 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1285  
OTHER INFORMATION: /note= "fusion protein composed of  
OTHER INFORMATION: enterokinase cleavable, His tagged  
OTHER INFORMATION: thioredoxin moiety and full length htrt"  
US-08-974-549A-600

Query Match 59.4%; Score 3505; DB 3; Length 1285;

Best Local Similarity 62.4%; Pred. No. 0;

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY	1	MTRAPRCPAVRSLLRSRYREVWPLATFVRRLLGPEGRRLLVQDPKIVTLVAQCLVCMHW	60
DB	154	MPEAPRCRAVRSLLRSRYREVWPLATFVRRLLGPEGRRLLVQDPKIVTLVAQCLVCMHW	213
QY	61	GSOPPPADLSFHOVSLKELVARVQRLCERNERNVLAFFELNEARGGPPMAFTSSVR	120
DB	214	DARPPPAAPSFQVSCLEKELVARVQRLCERNERNVLAFFELNEARGGPPMAFTSSVR	273
QY	121	SYLPNTVITLTVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA	180
DB	274	SYLPNTVITLTVSGAWMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA	333
QY	181	TTDIWPSVASRYRTPVGRNFTNLFQIQIKSSRQBPAPKLPALPSRGTKRHLSTSTS	240
DB	334	ATQARPPPHAS-GPARRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGASRS	384
QY	241	VPSAKZARCYVPVVEEGP-----HRQVLPTEGKSW-VPSPARSPENVTAOKDLSK	292
DB	385	LPLPKPRRGAPEPEPTVVGOGSWAHQRTGPDGFCVVSAR-----PAEATSL	439
QY	293	GKVSIDLSS-GSVCCKHFSSTLSPPRONAFQRP-FIETRHFLYSRGGQORLNPSF	350
DB	440	GALSGTRHSHPSVGRQHAGPPSTSRPPRPMPTPCPPVVAETKHPFLYSSGD-KEQLRPSF	498
QY	351	LLSNLOPNTGARRLVEIIFLGRSPTSGPLCRTHLSRYWQMRPLFQQLLVNHAECQY	410
DB	499	LLSSLRPSLTGARRLVETIFLGRSPTSGPLCRTHLSRYWQMRPLFQQLLVNHAECQY	558
QY	411	VRLLRSHCRFTANQOVTDAL-----NTSPHMLDLRLHSSPMQVY	452
DB	559	GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEDTPRRLVQLLRQHSPPWQVY	615
QY	453	GFRLACLCKVVSASLWGTNRNRRFFKLLKFIISLGKYKLSLOELMKMKVDECHLRS	512
DB	616	GFVRACLRLVPLGLWGSRRNRRFLNRTKFIISLGKAKLSLOELTWKMSVRDCAMLR	675
QY	513	SPGKDRVPAAEHRLRERILATFLFWMIDTVVQLRSFFYITESTFQKNLFFYRKSWS	572
DB	676	SPGVGCVPAAEHRLRERILATFLFWMIDTVVQLRSFFYITESTFQKNLFFYRKSWS	735
QY	573	KLOSIGVROHLEVRRLRELSQBEVRHQDTWLMPICRIRIPKPNGLRPIVNMYSMT	632
DB	736	KLOSIGIRQHLKTVQLRELSAEVRHQREARPALITSLRIFPKPDGLRPIVNMIDYVGA	795
QY	633	RALGRBKQAHFTQRLKTLFSLMNVERTKHPHMGSSVLGNDIVRTWRAFLVRALDQ	692
DB	796	RTFRREKRAELTSRKALFSLNVERARPPGLLAGSLGLDDIHEANRTFVLVRAQDP	855
QY	693	TPRMIFYADVTGAYDAIPQGLVEVANNMHSSTYICIRQYAVVRDSQGVHKSFR	752
DB	856	PELYFKVDVTGAYDTIPQDRLTEVIASIIK-PQNTYCVRRYAVVQAAHGHVRKAFKS	914
QY	753	QVTTLSLQPYMQOFLKHLQDSNLRNSVIEGOSISMNSSSSSLPFFELHRLHSVVK	812
DB	915	HVSTLTDLQPYMRQFVAHQET--SPLRDVAVVIEGSSLINEASSGLFVDFLRFMCHAVR	972





GenCore version 5.1.6  
 Copyright (c) 1993 - 2004 CompuGen Ltd.  
 protein - protein search, using sw model  
 n on: February 4, 2004, 14:11:55 ; Search time 46 Seconds  
 (without alignments)  
 5107.107 Million cell updates/sec  
 tle: US-09-042-460-2  
 rfect score: 5901  
 quence: 1 MTRAPRCRAVRLSRVRE.....TILKAAADPALSTDFQTILD 1122  
 bring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 arched: 801455 seqs, 209382283 residues  
 tal number of hits satisfying chosen parameters: 801455

nimum DB seq length: 0  
 ximum DB seq length: 2000000000  
 st-processing: Minimum Match 0\*  
 Maximum Match 100\*  
 Listing first 45 summaries

tabase : Published Applications AA:\*  
 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*  
 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pcp.\*  
 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*  
 4: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*  
 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
 7: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*  
 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*  
 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*  
 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*  
 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*  
 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*  
 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*  
 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	4751	80.5	1128	12	US-10-295-681-53
2	3505	59.4	1132	10	US-09-990-080-2
3	3505	59.4	1132	10	US-09-749-7288-31
4	3505	59.4	1132	10	US-09-843-676-225
5	3505	59.4	1132	10	US-09-953-052-2
6	3505	59.4	1132	12	US-10-295-681-57
7	3505	59.4	1132	15	US-10-053-758-225
8	3505	59.4	1132	15	US-10-208-243-2
9	3505	59.4	1132	15	US-10-054-295-225
10	3505	59.4	1132	15	US-10-054-611-225
11	3505	59.4	1132	15	US-10-054-611-225
12	3505	59.4	1132	15	US-10-044-692-2
13	3505	59.4	1132	15	US-10-044-539-2
14	3505	59.4	1154	15	US-10-044-692-323
15	3505	59.4	1154	15	US-10-044-539-323

16	3505	59.4	1189	15	US-10-044-692-325	Sequence 325, App
17	3505	59.4	1189	15	US-10-044-539-325	Sequence 325, App
18	3505	59.4	1200	15	US-10-044-692-324	Sequence 324, App
19	3505	59.4	1200	15	US-10-044-539-324	Sequence 324, App
20	3505	59.4	1285	15	US-10-044-692-314	Sequence 314, App
21	3505	59.4	1285	15	US-10-044-539-314	Sequence 314, App
22	3498	59.3	1132	12	US-10-385-882-2	Sequence 2, Appli
23	3496	59.2	1407	15	US-10-044-692-334	Sequence 334, App
24	3496	59.2	1407	15	US-10-044-539-334	Sequence 334, App
25	2492.5	42.2	1003	11	US-09-843-676-217	Sequence 217, App
26	2492.5	42.2	1003	11	US-09-438-486-217	Sequence 217, App
27	2492.5	42.2	1003	15	US-10-053-758-217	Sequence 217, App
28	2492.5	42.2	1003	15	US-10-054-295-217	Sequence 217, App
29	2492.5	42.2	1003	15	US-10-054-611-217	Sequence 217, App
30	2355	39.9	1131	12	US-10-295-681-52	Sequence 52, Appli
31	2195	37.2	807	15	US-10-044-692-5	Sequence 5, Appli
32	2195	37.2	807	15	US-10-044-539-5	Sequence 5, Appli
33	2132	36.1	622	15	US-10-294-778-12	Sequence 12, Appl
34	1825.5	30.9	500	12	US-10-282-960-81	Sequence 81, Appl
35	1686.5	28.6	364	12	US-10-304-095-41	Sequence 41, Appl
36	1511.5	25.6	438	15	US-10-294-778-10	Sequence 10, Appl
37	1506.5	25.5	564	10	US-09-843-676-101	Sequence 101, App
38	1506.5	25.5	564	10	US-09-766-253-101	Sequence 101, App
39	1506.5	25.5	564	11	US-09-438-486-101	Sequence 101, App
40	1506.5	25.5	564	15	US-10-053-758-101	Sequence 101, App
41	1506.5	25.5	564	15	US-10-054-295-101	Sequence 101, App
42	1506.5	25.5	564	15	US-10-054-611-101	Sequence 101, App
43	1497	25.4	437	15	US-10-294-778-2	Sequence 2, Appli
44	1249.5	21.2	364	12	US-10-304-095-40	Sequence 40, Appl
45	1105.5	18.7	538	15	US-10-044-692-316	Sequence 316, App

ALIGNMENTS

RESULT 1  
 US-10-295-681-53  
 ; Sequence 53, Application US/10295681  
 ; Publication No. US20030166270A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: E. Premkumar Reddy  
 ; APPLICANT: Sushil G. Rane  
 ; APPLICANT: Richard V. Mettous  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REVERSIBLY  
 ; INDUCING CONTINUAL GROWTH IN NORMAL CELLS  
 ; FILE REFERENCE: 6056-307  
 ; CURRENT APPLICATION NUMBER: US/10/295,681  
 ; CURRENT FILING DATE: 2002-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/334,760  
 ; PRIOR FILING DATE: 2001-11-15  
 ; NUMBER OF SEQ ID NOS: 69  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 53  
 ; LENGTH: 1128  
 ; TYPE: PRT  
 ; ORGANISM: Mesocricetus auratus  
 US-10-295-681-53

Query Match	80.5%	Score 4751;	DB 12;	Length 1128;
Best Local Similarity	80.1%	Pred. No. 0;		
Matches 304;	Conservative	90;	Mismatches 127;	Indels 8;
Gaps				4;
Qy	1	MTRAPRCRAVRLSRVREVPVPLATFVRLGEGRRLLVQGPDKIYRILVAQCLVCMHW	60	
Db	1	MTRAPRCRAVRLSRVREVPVPLATFVRLGEGRRLLVQGPDKIYRILVAQCLVCMHW	60	
Qy	61	GSQPPADLSFHVSSLSKELVARVQVRLCERNRNVLAFQFLLNEARGPPNMAFTSSVR	120	
Db	61	GSQPPADLSFHVSSLSKELVARVQVRLCERNRNVLAFQFLLNEARGPPNMAFTSSVR	120	
Qy	121	SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALYLLPPSCAYQVCGSPLYQICA	180	
Db	121	SYLPTNTVETLRVSGAWMLLSRVGDDLLVYLLAHCALYLLPPSCAYQVCGSPLYQICA	180	

181 TTDIWPVSASYPTRPVGRNFTNRLPQOIKSSQEQAPKPLALPSRGTQKHLSTSTS 240  
181 TAETWPSVSIYRPTREVGNGFTHLGSTRVNSHQEAWKPPPLPSREAKSLITNRS 240  
241 VPSAKKARCVFVRVEBGPQRVLPTPSGKSWPSPARSPVP--TAEDLSSGKQVSD 297  
241 VPPSKKARCDLAPLEKGPYQAVPTPSDKTWPNPAKSHAVPISTTKEDLSSGVKAPG 300  
298 LSLSGSVCKKHKSSTSLSPQONAPQLRPTFTRHLYSGDGOERLNPSFLLSNLQ 357  
301 LSRGSGVCYKHKPSSTLSQPLQONAPQLRPTFTRHLYSGDGOERLNPSFLLSNLQ 360  
358 NLTCARLVEIIFLGSRPRTSGPLCRTHLSRRYQWRPLFOCLLVNHAECYVLLRSH 417  
361 SLTGARLVELFLGMRPRTSGPLCGRRRLSKYQWRPLFOCLLVNHAECYVLLRSH 420  
418 CRRTANQVTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCVKVVSASLWGTNRER 476  
421 CRFRTAAHQVAGALNTTSPQRLMNLRLHSSPMQVYGFRLACLCVKVVSASLWGTNRER 480  
477 FPKNLKFIISLGKYSLSQELMWMKQVEDCHLRSRSPGKDRVPAEAHRLRERILATLF 536  
481 FFKVXKFIISLGKYSLSQELMWMKQVEDCHLRSRSPGKDRVPAEAHRLRERILATLF 540  
537 WLMQTYVYVQLLRSFFYITESTFQKRLFFYRKSVWSKLSIGVQRHLERVRRLSBEV 596  
541 WLMQTYVYVQLLRSFFYITESTFQKRLFFYRKSVWSKLSIGVQRHLERVRRLSBEV 600  
597 RHQDWTLMPICBRLRIPKPNGLRPTVNNMSYNGTRALGRKQAOHFTQRLTFSMLN 656  
601 RORQEAWFAMPICRLRIPKPNGLRPTVNNMSY-NGTRAFDKGQAOHFTQRLTFSMLN 659  
657 YERTKHPHMGSSVGLGNDIYRTWRAVFLAVRALDQTPRMVFKADVTGAYDAIPQGLV 716  
660 YELTKHNLGASVGLGNDIYRTWRAVFLAVRALDQTPRMVFKADVTGAYDAIPQGLV 719  
717 EVANMIRHSESTYCIQYAVVRDSQGVHKSFRQVTTLSDLQPYMGQFLLKHQSDA 776  
720 EVIANMIRHSPNSYCIHQYAVVRDSQGVHKSFRQVTTLSDLQPYMGQFLLKHQSDA 779  
777 SALNSVVIQSIISMNSSSLDFPLHPLRHSVVKIGRCYTCOCGIPQSGSLTLICS 836  
780 SALNSVVIQSIISMNSSSLDFPLHPLRHSVVKIGRCYTCOCGIPQSGSLTLICS 839  
837 LCFGDMENKLFAYQORDGLLRFVDDLLVTPHLDQAKTFLSTLVHGVPEYGCMLNLOXT 896  
840 LCFGDMENKLFAYQORDGLLRFVDDLLVTPHLDQAKTFLSTLVHGVPEYGCMLNLOXT 899  
897 VYNEPVEPTLGGAPQLPAHCLFPWCGLLDTPLEPCDYSGYACTSIKSLTFQSV 956  
900 VYNEPVDAGTLDGTAPHQLPAHCLFPWCGLLDTPLEPCDYSGYACTSIKSLTFQSV 959  
957 FKAGTKMKNLLSVLRKCHGLFELDQVNSLQVNCINIKIFLLOAYRPHACVLOLPDQ 1016  
960 FKAGTKMKNLLSVLRKCHGLFELDQVNSLQVNCINIKIFLLOAYRPHACVLOLPDQ 1019  
1017 RYRKNLTPFLGIIISQASCCYAILKVNQPGWTK-ASGSFPPEAAHMLCYQAFLLKLA 1073  
1020 HYRKNAPAFLLIIISQASCCYAILKVNQPGWTK-ASGSFPPEAAHMLCYQAFLLKLA 1079  
1074 AHSVYKCLGLGLPATAOKLCKRLPEATMTILKAAADPALSTDFQILD 1122  
1080 GHSVYKCLGLGLPATAOKLCKRLPEATMTILKAAADPALSTDFQILD 1128

## RESULT 2

3-09-990-080-2  
Sequence 2, Application US/09990080  
Patent No. US20020102686A1  
GENERAL INFORMATION:  
APPLICANT: Morin, Gregg B.  
APPLICANT: Genen Corporation

; TITLE OF INVENTION: Human Telomerase Catalytic Subunit Variants

; FILE REFERENCE: 018/258C

; CURRENT APPLICATION NUMBER: US/09/990,080

; PRIOR FILING DATE: 1998-08-03

; PRIOR APPLICATION NUMBER: US 09/052,864

; PRIOR FILING DATE: 1998-03-31

; NUMBER OF SEQ ID NOS: 21

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1132

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-990-080-2

Query Match 59.4%; Score 3505; DB 10; Length 1132;

Best Local Similarity 62.4%; Pred. No. 1.6e-316; Indels 52; Gaps 13;

Matches 719; Conservative 122; Mismatches 260;

Qy 1 MTRAPRCAPVRSLLRSRYREVMPLATFVRRLLGPEGRRLVQDPDKIYRTLVLAQCLVCMHW 60

Db 1 MPEAPRCAPVRSLLRSRYREVMPLATFVRRLLGPEGRRLVQDPDKIYRTLVLAQCLVCMHW 60

Qy 61 GSOPPPADLSFHOVSSLSKELVARVVRQRCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120

Db 61 DASPPPAAPSFQVQVCLKELVARVVRQRCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120

Qy 121 SYLNTVITLTVSGAMMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180

Db 121 SYLNTVITLTVSGAMMLLSRVGDDLLVLLAHCALYLLVPPSCAYQVCGSPLYQICA 180

Qy 181 TTDIWPVSASYPTRPVGRNFTNRLPQOIKSSQEQAPKPLALPSRGTQKHLSTSTS 240

Db 181 ATQARPPPHAS-GPERRLG-----CERAMNHSVREAGVPLGLPAPGARRRGSGASRS 231

Qy 241 VPSAKKARCVFVRVEBGP-----HRQVLTSPGKSW-VSPSPARSPEVETAEDLSSK 292

Db 232 LPLPKPRRGAPEBERTPVGGSWAHFGRTRGSDRGFCVVSFAR-----PABEATSL 286

Qy 293 GKVDLSLS-GSVCKKHKSSTLSLSPQONAPQLRPTFTRHLYSGDGOERLNPSF 350

Db 287 GALSGRTHSHSPVGRGHAGPSTSRPPRMDTPCPVYVETKGFILYSSGD-KEQLRPSF 345

Qy 351 LLSNLQPNLTGARRLVEIIFLGSRPRTSGPLCRTHLSRRYQWRPLFOCLLVNHAECY 410

Db 346 LLSLSPSTGTARRLVETIIFLGSRPMPPTPRPLPQRYQWRPLFLELLGHAQCPY 405

Qy 411 VRLRHGCRFRFANQVTDAL-----NTSPHMLDLRLHSSPMQVY 452

Db 406 GVLLKTHCLRAA---VTPAAGVCAREKPGSVAAPBEEDTPRLVQLLRQHSSPMQVY 462

Qy 453 GFLRACLCKVVGASLWGTNRHNRFRFKNLKKFISLGKYSLSQELMWMKQVEDCHLRS 512

Db 463 GFYRACLRVLFPGLWGSRRNRFRFRNTKPFISLGKYSLSQELMWMKQVEDCHLRS 522

Qy 513 SPCKDRVPAEAHRLRERILATFLWMDTYVYVQLLRSFFYITESTFQKRLFFYRKSVWS 572

Db 523 SPVGVCPVPAEAHRLRERILATFLWMDTYVYVQLLRSFFYITESTFQKRLFFYRKSVWS 582

Qy 573 KLSQIGVGRHLERVRRLSBEVHHQDWTLMAMPICRLRIPKPNGLRPTVNNMSYNGT 632

Db 583 KLSQIGVGRHLERVRRLSBEVHHQDWTLMAMPICRLRIPKPNGLRPTVNNMSYNGT 642

Qy 633 RALGRKQAOHFTQRLTFLWMDTYVYVQLLRSFFYITESTFQKRLFFYRKSVWS 692

Db 643 RTRRKRERERTSRVYALFVINTERRARPGGLGSLGGLDDIHRWRTFVLVRADQP 702

Qy 693 TPRMYFVKADVTGAYDAIPQGLVVEVAMNIRHSESTYCIQYAVVRDSQGVHKSFR 752

Db 703 PPELYFVKVDVTGAYDTIPQDLTEVIAIISK-PQNTYCVARYAVWQAAHGHVKAFKS 761

Qy 753 QVTTLSDLQPYMGQFLLKHQSDASALNSVVIQSIISMNSSSLDFPLHPLRHSVVK 812

Db 762 HVSTLTLDQPYMRQRFVAHLQET--SPLRDVAVVIEOSSLSNASSGLFVDFLRFMCHAVR 819

813 IGRDCTVQCGIPQSSSLTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 872  
 820 IRGKSVQCGIPQSSSLTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 879  
 873 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEGTLGGAPQLPAHCLFPWCGLLLDTOT 932  
 880 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEGTLGGAPQLPAHCLFPWCGLLLDTOT 939  
 933 LEVFCDSYGAQTSIKTSUTFOSVFKAGKTMNKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASLTFRNGFKAGRMNRKLFGLVLRKCHGLFLDLQVNSLQTVCI 999  
 993 NIYKIFLLQAYRFHACVQLPQDQVRKNTLFFLGISSQASCCVAILKVNPGWMLKAS 1052  
 1000 NIYKIFLLQAYRFHACVQLPQDQVRKNTLFFLGISSQASCCVAILKVNPGWMLKAS 1059  
 1053 GS---FPPEAAHWLCYQAFLLKLAHSHVYIKCLGLGLRTPAQKLLCRKLPBATMTILKAAA 1109  
 1060 GAAGPLPSEAVQWLCHQAFLLKLTFRHTVYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQTLTD 1122  
 1120 NPALPSDFKTILD 1132

SULT 3  
 -09-749-728B-31  
 Sequence 31, Application US/09749728B  
 Patent No. US20020142457A1  
 GENERAL INFORMATION:  
 APPLICANT: Umezawa, Akhihiro  
 APPLICANT: Hata, Jun-ichi  
 APPLICANT: Fukuda, Keiichi  
 APPLICANT: Osawa, Satoshi  
 APPLICANT: Sakurada, Kazuhiro  
 APPLICANT: Gojo, Satoshi  
 APPLICANT: Yamada, Yoji  
 TITLE OF INVENTION: THE CELL HAVING THE POTENTIALITY OF DIFFERENTIATION INTO CARDIOMY  
 FILE REFERENCE: 00766.000043  
 CURRENT APPLICATION NUMBER: US/09/749,728B  
 CURRENT FILING DATE: 2001-09-17  
 PRIOR APPLICATION NUMBER: H11-372826  
 PRIOR FILING DATE: 1999-12-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-01148  
 PRIOR FILING DATE: 2000-02-28  
 PRIOR APPLICATION NUMBER: PCT-JP00-07741  
 PRIOR FILING DATE: 2000-11-02  
 NUMBER OF SEQ ID NOS: 80  
 SOFTWARE: Patentin ver.2.0  
 SEQ ID NO 31  
 LENGTH: 1132  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 -09-749-728B-31

Query Match 59.4%; Score 3505; DB 10; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
 1 MTRAPRCFVRSLLRSYREVWPLATFVRRLGPEGRVLVQGDPKIYRTLVAQCLVCMHW 60  
 1 MPAPRCFVRSLLRSYREVWPLATFVRRLGPEGRVLVQGDPAFRLVAQCLVCMHW 60  
 61 GSOPFPADLSFHQVSSIKELVAVVQRLCERNVNLAFGPELLNEARGGPMFTSSVR 120  
 61 DARPPPAAPSPRQVSLKELVAVVQRLCERNVNLAFGPELLNEARGGPMFTSSVR 120  
 121 SYLPTNTVETLRVSGAWMLLSVGGDDLLVLLAHCALVLLVPPSCAVQVCGSLYQICA 180  
 121 SYLPTNTVETLRVSGAWMLLSVGGDDLLVLLAHCALVLLVPPSCAVQVCGSLYQICA 180  
 181 TTIDTFVSVSAYRPTFRVGRNFTNLRFLQIKSSSSQEAAPKPLALPSRGTKRHLSTSTS 240

181 ATQARPPPHAS-GRRRRLG-----CERAWNSHREAGVPLGLFAPGARRRGGASRS 231  
 241 VPSAKKARCYVPRVEGCP-----HROVLTPSGKSM-VPSPARSPVEPTAEKOLSSX 292  
 232 LPLPRPRRGAAPERTFVQGGSWAHPLGRTRGSDRGFCVSPAR-----PAEATSLSE 286  
 293 GKVSLSLS-GSVCKKHPSTSTLSPPQNAFOLRP-FIETRHFLYSRGGGOERLNPSF 350  
 287 GALSSTRSHSVGRQHAGPPSTSRPPRMDTCCPPVYATKHFLLYSYSGD-KEQURPSF 345  
 351 LLSNLQPNLTGARRLVEIIFLGSRPSTSGPLCRTHLSRRYWMQRPFLQOLLVNHAEQY 410  
 346 LLSLRPSLTGARRLVEIIFLGSRPMPGTPRRLPRLPQRYWMQRPFLFLELLNHAQCPY 405  
 411 VLLASHCRFRFANQVTDAL-----NTSPHLMDLILRLHSSPMQVY 452  
 406 GVLLKTHCPLRAA---VTPAAGVCAREKPOGSAAPBEEDTPRRLVQLLRQHSSPMQVY 462  
 453 GFLRACLCKVYSASLWGRNRRRPFKNLKKFISLGRYKGLSLQELMWMKVEDCHWLS 512  
 463 GFVRACLRLVPPGLWGRNRRRFLNTKKFISLGRYKGLSLQELMWMKVEDCHWLS 522  
 513 SPGKDRVPAABERLPERILATFLWMDTYVYVQLLRSFFYITESTFQKNLFFYRKSWS 572  
 523 SPGVGVZAAERLREELAKFLHLMVYVVELLSFFVYVTTETFOKNLFFYRKSWS 582  
 573 KLOSGIVRCHLRLRELSCOEVRHHDCTWLAMPICRLRFIPKPNGLRPIVNSMSGT 632  
 583 KLOSGIVRCHLRLRELSCOEVRHHDCTWLAMPICRLRFIPKPNGLRPIVNSMSGT 642  
 633 RALGRKQAOHTORLKLTFSLMNYERTKPHLMGSSVLGMNDIYRTWRAFVLRALDQ 692  
 643 RTFREKRAERLTSRVKALFSLVNYERARRPGLLGASVGLGDDIHRAWRTFVLVRAQDP 702  
 693 TPRMYVADVTGAYDAIPQKLVVAVNMIHSESTYCIQYAVVTRDSCQVHKSFR 752  
 703 PPELYFVKVDVTGAYDTIPQDLTEVIAIHK-PONTYCVRRYAVVQAAHGHVKAFFS 761  
 753 QVTTLSLDQPMYQGLKHLQSDASALNSVVIQSSIMNESSSLEDFLHFLRHSHVWK 812  
 762 HVTSLTDLPYMRQFVAHQET--SPLDVAVIEQSSSLNEASSGLFDVFLRFMCHAVR 819  
 813 IGRDCTVQCGIPQSSSLTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 872  
 820 IRGKSVQCGIPQSSSLTLLCSLCFQDMENKLPFAEVQORGLLLRFVDDFLVTPHLDQ 879  
 873 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEGTLGGAPQLPAHCLFPWCGLLLDTOT 932  
 880 AKTFELSTLVHGVPEYGCMLNOKTVNRPVEGTLGGAPQLPAHCLFPWCGLLLDTOT 939  
 933 LEVFCDSYGAQTSIKTSUTFOSVFKAGKTMNKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSYARTSIRASLTFRNGFKAGRMNRKLFGLVLRKCHGLFLDLQVNSLQTVCI 999  
 993 NIYKIFLLQAYRFHACVQLPQDQVRKNTLFFLGISSQASCCVAILKVNPGWMLKAS 1052  
 1000 NIYKIFLLQAYRFHACVQLPQDQVRKNTLFFLGISSQASCCVAILKVNPGWMLKAS 1059  
 1053 GS---FPPEAAHWLCYQAFLLKLAHSHVYIKCLGLGLRTPAQKLLCRKLPBATMTILKAAA 1109  
 1060 GAAGPLPSEAVQWLCHQAFLLKLTFRHTVYVPLGLSLRTAQQLSRKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQTLTD 1122  
 1120 NPALPSDFKTILD 1132

RESULT 4  
 US-09-843-676-225  
 ; Sequence 225, Application US/09843676  
 ; Patent No. US20020164786A1  
 ; GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
 Lingner, Joachim  
 Nakamura, Toru  
 Chapman, Karen B.  
 Morin, Gregg B.  
 Harley, Calvin  
 Andrews, William H.  
 TITLE OF INVENTION: No. US20020164786A1el Telomerase  
 NUMBER OF SEQUENCES: 225  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, 8th Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94111

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/843,676  
 FILING DATE: 26-Apr-2001  
 CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/854,050  
 FILING DATE: 09-MAY-1997  
 APPLICATION NUMBER: US 08/846,017  
 FILING DATE: 25-APR-1997  
 APPLICATION NUMBER: US 08/844,419  
 FILING DATE: 18-APR-1997  
 APPLICATION NUMBER: US 08/724,643  
 FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:  
 NAME: Apple, Randolph T.  
 REGISTRATION NUMBER: 36,429  
 REFERENCE/DOCKET NUMBER: 015389-002930US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 576-0200  
 TELEFAX: (415) 576-0300  
 INFORMATION FOR SEQ ID NO: 225:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 225:

S-09-843-676-225

Query Match 59.4%; Score 3505; DB 10; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.66-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

Y 1 MTRAPRCAPASLLRSRYREVPLATFVRRLGPEGRLVQGPDKYRVLVAQCLVCMHW 60  
 b 1 MTRAPRCAPASLLRSRYREVPLATFVRRLGPEGRLVQGPDKYRVLVAQCLVCMHW 60

Y 61 SQPPADLSHQVSSLSKELVARVQRLCERNERNVLAFCFELLNARGGPPMAFTSSVR 120  
 b 61 DARPPPAAPSPRQVSCLELVARVQRLCERNERNVLAFCFELLNARGGPPMAFTSSVR 120

Y 121 SYLPTNTVETLRVSGAWMLLSRYGDDLLVYLLAHCAALYLLVPPSCAYQVCGSPLYQICA 180  
 b 121 SYLPTNTVETLRVSGAWMLLSRYGDDLLVYLLAHCAALYLLVPPSCAYQVCGSPLYQICA 180

Y 181 TTDIWPVSASRYRTPVGRNFTNLRFLQKQSSRQAPKALPSRGKRLHLSLTS 240  
 b 181 ATQARPPPHAS-GRPRRLG-----CERAWNSHVEAGVPLCLPAPGARRRGGASRS 231

Y 241 VPSAKKARCYVPVPRVEGP-----HRQVLPTSGKSW-VPSPARSPVPTAEKLSK 292  
 b 241 VPSAKKARCYVPVPRVEGP-----HRQVLPTSGKSW-VPSPARSPVPTAEKLSK 292

Y 232 LPLPKRRRGAPEPERTVPGGSGWAHPGRTRGSDRGFCVVSFAR-----PAEATSLE 286  
 b 232 LPLPKRRRGAPEPERTVPGGSGWAHPGRTRGSDRGFCVVSFAR-----PAEATSLE 286

QY 293 GKVDLSLS-GSVCKKHPSPSTLSLSPRQNAFOLRP-FIETRFYLSRGGQERLNPSF 350  
 Db 287 GALSOTRSHSPSVGQHGAGPSTSRPPRWDTPCPVPYATKHFYSSGD-KEQLRPSF 345  
 QY 351 LLSNLQPNLTGARRLVEIFLGSRPRTSGPLCRHLSRRYQWRPLFOQLLVNHAFCQY 410  
 Db 346 LLSLRPSLTGARRLVEIFLGSRPRTSGPLCRHLSRRYQWRPLFOQLLVNHAFCQY 405  
 QY 411 VLLRSHCRFRTANQOVTDAL-----NTSPHMLDLARLHSSPQVY 452  
 Db 406 GYLLKTHCFELRAA---VTPAAGVCAREKPGSVAPBEEDTPRRLVQLLRQHSPPQVY 462  
 QY 453 GFLRACLCKVVSASLWGTNRHNERFFKNLKKFISLGYKGLSLQELMWKMKVEDCHWLS 512  
 Db 463 GFVRACLRLVPPGLWGSRRHNERFLRNTKFI SLGXHAKLSLQELTWKMSVRDCAWLR 522  
 QY 513 SPKORVPAAEHRLRERILATFLWMDTYVQLRSPFFYITESTFOKNRLFYFKKSVMS 572  
 Db 523 SPQVGCVPAAEHLREELAKFLHWSVYVVELLRSPFFYITESTFOKNRLFYFKKSVMS 582  
 QY 573 KLSIGVROHLERVLRELSEVSHHODTWLAMPICRLRFIPKPNGLRPTVNMSSMGT 632  
 Db 583 KLSIGVROHLERVLRELSEVSHHODTWLAMPICRLRFIPKPNGLRPTVNMSSMGT 642  
 QY 633 RALGRKQAOHFTQRLKTLFSLMNYERTKPHLMGSSVLGMNDIYRTWRAPVLVRLDQ 692  
 Db 643 RTFRREKRAERLTSRKALFSLVNYERARRPCLLGASVLGLDDIHRARWTFVLVRADP 702  
 QY 693 TPRMEVKADVTGAYDAIPQKLVWVANMIRHSESTYCIROYAVVRRDSOGVHKSFRR 752  
 Db 703 PPELYFVKVDVTGADTTPQDLTEVIAIILK-PQNTYCVRYAVVQKAAHGHVYKAFKS 761  
 QY 753 QVTLSDLPYMGQFLKHLQSDASALRNSVVI EOSISMNESSSLEDFLHFLRHSVVK 812  
 Db 762 HVSITLTLQPYMRFVHLEQET--SPLRDVAVIEGSSSLNEASSGLEDFVFLRFMCHAVR 819  
 QY 813 IGRDCTYQCGIPQSSSLTLLCSLCFGDMENKLPFAVQORDGLLRFVDDFLVTPHLQ 872  
 Db 820 IRGKSYVQCGIPQSSSLTLLCSLCFGDMENKLPFAVQORDGLLRFVDDFLVTPHLQ 879  
 QY 873 AKTFSLTLVHGVPYEGCMINLQKTVVNFPEFTLGGAAAPYQLPAHCLFPWCGLLDTQT 932  
 Db 880 AKTFSLTLVHGVPYEGCMINLQKTVVNFPEFTLGGAAAPYQLPAHCLFPWCGLLDTQT 939  
 QY 933 LEVFCDSYGAOTSITSLTFSQVFKAGTKWKNLLSVLRKCHGLFLDLQVNSLTQVCI 992  
 Db 940 LEVQSDYSYARTSIRASLTFRNGFKAGNMRERKLFGLVRLKCHSLFLDLQVNSLTQVCI 999  
 QY 993 NIYKIFLLQAYRPHACVIOLPFDQVRKLTFFLGIISQASCCVAILKVNPGMTLKAS 1052  
 Db 1000 NIYKIFLLQAYRPHACVIOLPFDQVRKLTFFLGIISQASCCVAILKVNPGMTLKAS 1059  
 QY 1053 GS---PFPBAHWLCYQAFLLKLAHSVIYKCLIGPLTAQKLCRKLPEATMILKAA 1109  
 Db 1060 GAAGPLPSEAVQWLCQAFLKLLTRHRTVYVPLGSLRTAQQLSRKLPGTTLTAAEA 1119  
 QY 1110 DPALSTDFQITLD 1122  
 Db 1120 NPALPSDFKITLD 1132

RESULT 5  
 US-09-953-052-2  
 ; Sequence 2, Application US/09953052  
 ; Patent No. US20020173476A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cech, Thomas R.  
 ; Lingner, Joachim  
 ; Nakamura, Toru  
 ; Chapman, Karen B.  
 ; Morin, Gregg B.  
 ; Harley, Calvin B.

Andrews, William H.  
TITLE OF INVENTION: Antisense Compositions for Detecting and  
Inhibiting Telomerase Reverse Transcriptase

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICANT: Townsend and Crew LLP

APPLICATION NUMBER: US/09/953,052

FILING DATE: 14-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/052,919

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/911,312

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/912,951

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/915,503

FILING DATE: 14-AUG-1997

APPLICATION NUMBER: US 08/974,549

FILING DATE: 19-NOV-1997

APPLICATION NUMBER: US 08/974,584

FILING DATE: 19-NOV-1997

APPLICATION NUMBER: WO PCT/US97/17618

FILING DATE: 01-OCT-1997

APPLICATION NUMBER: WO PCT/US97/17885

FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:

NAME: Parent, Annette S.

REGISTRATION NUMBER: 42,058

REFERENCE/DOCKET NUMBER: 015389-00360005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1132 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

-09-953-052-2

Query Match

Best Local Similarity

Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCRVAVRSLRSHYREVMPLATFVRLGPEGRRLVQDPDKIYRTILVAQCLVCMHW 60

1 MTRAPCRVAVRSLRSHYREVMPLATFVRLGPEGRRLVQDPDKIYRTILVAQCLVCMHW 60

61 GSQPPPADLSFHVQSLKELVARVQVORLERNVLAFGPELLNEARGGPPNAFTSSVR 120

61 DARPPPAAPSFQVSLKELVARVQVORLERNVLAFGPELLNEARGGPPNAFTSSVR 120

121 SYLNTVETLRVSGAWMLLSRVGDDLLVYLALHACALYLLVPPSCAYQVCGSPLYQICA 180  
121 SYLNTVETLRVSGAWMLLSRVGDDLLVYLALHACALYLLVPPSCAYQVCGSPLYQICA 180  
181 TTDIWPVSASRYRTPVGRNFTNLRFLQOIKSSSQEAPKPLALPSRGTKRHLSTLSTS 240  
181 ATQARPPPHAS-GRPRRLG-----CERAWNHSYREAGVPLGLPAPGARRRGGSASRS 231  
241 VPSAKKARCYVPVREVEGP-----HRQVLPTPSKSM--VPSPARSPVPTAKOLSKS 292  
232 LPLPKRPRRGAAPFERTPVQGSWAHPGTRGSDRGFCVGVSPAR-----PAEEATSL 286  
293 GKVSDELISL-GSVCKKHKPSSTLSLSPRONAQRLRP-FIETRHFLYSRGDGOERLNPSF 350  
287 GALSGRHSHPSVGRQHAGPPSTSRPPRPWDTPCPVVAETKHFYSSGD-KEQRLRPF 345  
351 LLSNLQPNLTGARELVEIIFLGRPRPTSGPLCTHRLSRRYWQMRPLFOOLLVNHAEQY 410  
346 LLSLRPSLTGARLVEIIFLGRPRPTSGPLCTHRLSRRYWQMRPLFOOLLVNHAEQY 405  
411 VLLRSRCPRTANQVTDAL-----NTSPHMLMDLRLHSSPWQY 452  
406 GVLLKTHCPURAA---VTPAGVCAREKPGSGVAAPEDDTPRRLVQLLRQHSWPQY 462  
453 GFLRACLCCKVVSASLWGRHNRERFFKNLKKFISLQKYGKLSLQELMWKVKVEDCHWLS 512  
463 GFVRACLRLRVLPPGLWGSRNERRFLNTKFFISLQKHAKLSLQELTWKMSVRDCAWLR 522  
513 SPGKDRVPAAEHRLRERILATFLWMDTVVOLLRSFFYIIESTFKNRLPFYKSVMS 572  
523 SPGVGCVPAAEHRLRERILAKFLWMSVVVVELLSFFYVTTETTFQKRLPFYKSVMS 582  
573 KLSQIGVROHLRVRRLRELSOEVRHHDQTLWAMPICRLRFIPKPNGLRPIVNMYSMT 632  
583 KLSQIGVROHLRVRRLRELSOEVRHHDQTLWAMPICRLRFIPKPNGLRPIVNMYSMT 642  
633 RALGRKQAOHFTORLKTFLSMLNERTKHPHMGSVLGMNDIYRTWRAFLVRVRLDQ 692  
643 RTFRREKAERLTSRVKALFVNLNTERARPPGLGASVLGLDDIHEAWRTFVLRVRAQDP 702  
693 TPRVYFKADVTGAYDAI POGKLVVVVNMIRHSESTYCIQAVVVRDSQGOVHKSFR 752  
703 PPVLYFKVADVTGAYDTI PDLRTVETASIK-PQNTYCVRRYAVVQKAAHGHVRKAFKS 761  
753 QVTLSDLPYMGQFLKHLQSDASALRNSVIEQSI SMNNESSLSLDFPLHFLRHSVYK 812  
762 HVSTLTDLPYMRQFAHLQET--SPLRDAVIEQSSSLNEASSGLDFVFLRFMCHAVR 819  
813 IGDRCYTCQCGIPGSSSLTLLCSLFCGDMENKLPFAVQDGLLRFVDDFLVTPHLDQ 872  
820 IRGKSYVQCGIPGSSSLTLLCSLFCGDMENKLPFAVQDGLLRFVDDFLVTPHLDQ 879  
873 AKTFLSTLVHGVPEYGCWMLNLTQVNVFPVPEPTLGGAAPIQLPAHCLFPWCGLLDQ 932  
880 AKTFLSTLVHGVPEYGCWMLNLTQVNVFPVPEPTLGGAAPIQLPAHCLFPWCGLLDQ 939  
933 LEVPCDYSVTAQTSIKTSLTPQSVFAGKTMKRLSVLRKCHGLFLLDQVNSLQTVCI 992  
940 LEVQSDYSVTAQTSIKTSLTPQSVFAGKTMKRLSVLRKCHGLFLLDQVNSLQTVCI 999  
993 NIYKIFLLQAVRFHACVQLPFDQVRKRLTFFLGIISQOASCCVAILKVKQVPMGLKAS 1052  
1000 NIYKIFLLQAVRFHACVQLPFDQVRKRLTFFLGIISQOASCCVAILKVKQVPMGLKAS 1059  
1053 GS---FPPEAHMICYQAFLKLAASHVYKLLGGLPRTAQKLLCKLPEATWTILKAAA 1109  
1060 GAAGFLPSEAVOMLCHQAFLLKLTTRHRTVYVPLGLSLRTAQKLLCKLPEATWTILKAAA 1119  
1110 DPALSTDFQILD 1122  
1120 NPALSPDFKILD 1132

SULT 6  
-10-295-681-57  
Sequence 57, Application US/10295681  
Publication No. US20030166270A1  
GENERAL INFORMATION:  
APPLICANT: E. Prekumar Reddy  
APPLICANT: Sushil G. Kane  
APPLICANT: Richard V. Mettus  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REVERSIBLY  
INDUCING CONTINUAL GROWTH IN NORMAL CELLS  
FILE REFERENCE: 6056-307  
CURRENT APPLICATION NUMBER: US/10/295,681  
CURRENT FILING DATE: 2002-11-15  
PRIOR APPLICATION NUMBER: US 60/334,760  
PRIOR FILING DATE: 2001-11-15  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 57  
LENGTH: 1132  
TYPE: PRT  
ORGANISM: Homo sapiens  
-10-295-681-57  
Query Match 59.4%; Score 3505; DB 12; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
/ 1 MTRAPRCNVRSLRSRYREVWPLATFVRLGPEGRLVQPGDKPIRYTLVAQCLVCMHW 60  
/ 1 MPRAPRCNVRSLRSRYREVWPLATFVRLGPEGRLVQPGDKPIRYTLVAQCLVCMHW 60  
/ 61 GSQPPADLSHFVSSLSKELVARVQVRLCERNERNVLAFGFELNENARGGPPMAFTSSVR 120  
/ 61 DARPPAPSPRQVSCUKELVARVQLQELCERGAKNVLAFGFALLDARGGPPMAFTSSVR 120  
/ 121 SYLNTVITLVRSGAMWLLSRVDDLLVLAHALYLLVPPSCAYOVCGSPYQICA 180  
/ 121 SYLNTVITDARGSGAMWLLSRVDDLLVLAHALYLLVPPSCAYOVCGSPYQICA 180  
/ 181 TTDIPWSVASYRTPRPGVGRNFTNRLFQIKSSRCQEAPEKPLALPSRGTKHLSLTS 240  
/ 181 ATQARPPHSA-GRPRIG-----CERAWNSVREAGVFLGLPAPARRGGSASES 231  
/ 241 VPSAKARCPVPVREBGP-----HROVLTPTSGKSM-VPSAPARSPEVPTAEKDLGSK 292  
/ 232 LPLPKRPRGAAPERTPTVQGGWAHPGTRGDSRDFGFCVWSPAR-----PABEATSLE 286  
/ 293 GKVSLSLS-CSVCKKPSSTLSLSPRQNAFLRP-FIETBHLYSRGDQOERLNPSF 350  
/ 287 GALSOTRSHSPVGRHAGPPTSRPRPDWDTCPPEVYAEKHFLLYSGD-KEQLRPSF 345  
/ 351 LLSNLQPNLTGARRLVETIFLGSRRPRTSGPLCRTHRLSRRYQWRPLFQOLLVNAECQY 410  
/ 346 LLSLSRPSLTGARRLVETIFLGSRRPRTSGPLCRTHRLSRRYQWRPLFQOLLVNAECQY 405  
/ 411 VALLRSHCRFTANQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
/ 406 GYLLKTHCPLEAA---VTPAAGVCAKPKQSGVAAPPEEDTDPRLVQLLRQHSPPWQY 462  
/ 453 GFLRACLCKVYSASLWGRNRRERFFKNKKFISLGYKLSQELMWKMKVDDCHWLRS 512  
/ 463 GFLRACLRLVPPGLWGRNRRERFFKNKKFISLGYKLSQELMWKMKVDDCHWLRS 522  
/ 513 SPGRDVRPAABHRLRERLATFLFWMNDTVVQLRSFFVITESTQKRLFFYRKSVM 572  
/ 523 SPGVGCVPAABHRLRERLATFLFWMNDTVVQLRSFFVITESTQKRLFFYRKSVM 582  
/ 573 KLOSGVCHLRLVRLSRLSOEVRHODTWLAMPICRLRFIPKNGRLFVNMVSMGT 632  
/ 583 KLOSGVCHLRLVRLSRLSOEVRHODTWLAMPICRLRFIPKNGRLFVNMVSMGT 642  
/ 633 RALGRKQAOHFTORLKLFLGMLNRYETKPHLMGSSVLGMNDIYRTWRAFLVRLALDQ 692

643 RTFREKRAERLTSRVKALFSLVNERARRPCLLGASVGLDDIHRARWTFVLVRQAQDP 702  
693 TPRMYFVKADVTGAYDAIPQKLVVWAMIRHSESTYCIQYAVVRRDSQGVHKSFR 752  
703 PPFLYFVKVDVTGAYDTIPQRLTEVIASIIK-PQNTYCVRRYAVVQAAHGVKAFKS 761  
753 QVTTLSLDLPYMGQELKHLQSDASALENSVVISOS:SMNESSSSSLDFELFLRHSVVK 812  
762 HVTSTLD:QPYNRQFVAHQET--SPLDAAVVIQSSSLNEASSGLFDVFLRFMCHAVR 819  
813 IGDRCYTCQGIPOGSSISLTLCSLCFCGDMENKLFVAVQDRLRLRVDDPFLVTPHLQ 872  
820 IRGKSVQCGIPQGSISLTLCSLCYCGDMENKLFAGIRRDGLLLVDDPFLVTPHLTH 879  
873 AKTFSTLVHGVPEYGCMMINLOKTVNPEPVGTLGGAAPYQLPAHCLFPMWGLLDTOT 932  
880 AKTFSTLVHGVPEYGCMMINLOKTVNPEPVGTLGGAAPYQLPAHCLFPMWGLLDTOT 939  
933 LEVPCDYSYAOISITSLTFSQVFKAGTKMENKLLSVLRKCHGLEFLDLQVNSLOTVCI 992  
940 LEVQSDYSYARTSIRASITFNRGFKAGNWRKLFGLVRLKCHSLFLLDLQVNSLOTVCI 999  
993 NIYKIFLQAYRFHACVQLPDPQVRKXNLTFFIGITISSQASCCYAILKVNPMOTLKAS 1052  
1000 NIYKIFLQAYRFHACVQLPDPQVRKXNLTFFIGITISSQASCCYAILKVNPMOTLKAS 1059  
1053 GS---PPPAHAWLCVQAFLLKLAHSAHVIYKCLLPRLTAQKLCRKLPEATMTILKAAA 1109  
1060 GAAGPLPSEAVONLCHQAFLLKLTHERVTVYVPLGLSLTAQTQLSRKLPGTTLTALBAAA 1119  
1110 DPALSTDFOTILD 1122  
1120 NPALPSDFKTILD 1132

RESULT 7  
US-10-053-758-225  
; Sequence 225, Application US/10053758  
; Publication No. US20030032075A1  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; Linsinger, Joachim  
; Nakamura, Toru  
; Chapman, Karen B.  
; Morin, Gregg B.  
; Harley, Calvin  
; Andrews, William H.  
; TITLE OF INVENTION: No. US20030032075A1el Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/053,758  
; FILING DATE: 18-Jan-2002  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/854,050  
; FILING DATE: 09-MAY-1997  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
-10-053-758-225

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
1 MTRAPRCRAVRSILRSRYREVWPLATFVRRLLGPEGRRRLVQCPDKIYRTLVAAQCLVCMHW 60  
1 MPAPRCRAVRSILRSRYREVWPLATFVRRLLGPEGRRRLVQCPDKIYRTLVAAQCLVCPW 60  
61 GSOPPPADLSFQVSSLSKELVARVVRQICERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
61 DARPPAPAFSFRQVSKELVARVVRQICERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
121 SYLNPNTVETLRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLYQICA 180  
121 SYLNPNTVETLRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLYQICA 180  
181 TTDIMPVSASVYRTPVGRNTNLRFTQQIKSSRQEAAPKLPALPSGTRKHLSLSTGS 240  
181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGSASRS 231  
241 VPSAKKARCVPRVVEEGP-----HRQVLTPTSGKSW-VPSPARSPVPTAEKDLSSK 292  
232 LPLPKPRGRGAPEPRTVPGGSAHFGTRTGPDSRGFCVVSAPR-----PAEATSL 286  
293 GKVDLSLS-GSYCCCKHPSSTLSPPRQNAFLRP-FIETRLFLXSRGQGRRLNPSF 350  
287 GALSGRHSHPSVGRQHHAGPSTSRPPRPMDTPCPVVAETKHFLYSSGD-KQLRPSF 345  
351 LLSNLOPNTGARRLVEIIFLSRPTSGPLCRHLSRRYQWRPLPQQLLVNHAECQY 410  
346 LLSLSLPSLTGARRLVETIFLGSRRPMPGTPRRLPLPQRYQWRPLPQQLLVNHAECQY 405  
411 VRLLRSHCRFRFTANQOVTDAL-----NTSPHMLDLLHSSPMQVY 452  
406 GVLLKTHCPLRAA---VTPAAGVCAKEKPGQSVAAPEEEDTPRLVQLLRQHSPPQVY 462  
453 GFURACLCKVVASLWGTNRHRRFKNLKKPISLGKYGKLSLOELMMKMKVEDCHMLRS 512  
463 GFVRACLRLVPPGLWGSRRHRRFKNLKKPISLGKYGKLSLOELMMKMKVEDCHMLRS 522  
513 SPQKDVPAEHLRLERILATLFLWMDTVVQVLLRSFFYTESTFQKRLFFFRKSVWS 572  
523 SPQGVCPVPAEHLRLERILATLFLWMDTVVQVLLRSFFYTESTFQKRLFFFRKSVWS 582  
573 KLOSIGVGRVQLRLERILATLFLWMDTVVQVLLRSFFYTESTFQKRLFFFRKSVWS 632  
583 KLOSIGVGRVQLRLERILATLFLWMDTVVQVLLRSFFYTESTFQKRLFFFRKSVWS 642  
633 RALGREGKQAHETQRLKTLFSLMNTYRKHPLHMGSSVLGNDIYRTVRAVLVRLDQ 692  
643 RTRPRKRAERTSRVKALFSLVNTYRKHPLHMGSSVLGNDIYRTVRAVLVRLDQ 702  
693 TRPMYFVKADVTGAYDAIPQGLKVEVAVNMRHSESTYCIQYAVVRDSDGQVHKSFRR 752

Db 703 PPELYFVKVDVTGAYDVTIPQDRLTEVTASIIK-PQNTYCVRRYAVVQKAAHGHVRKAFKS 761  
QY 753 QVTTLSLDLPYMGOFKLHQLQSDASALRNSVVEIQSISMNNESSSLFDFFLHFLRHSVVK 812  
Db 762 HVSITLTDLPYMRQFVAHQET--SPURDAVVEIQSSSLNEASSGLFDVFLRFMCHAVR 819  
QY 813 IGDRCYTQCGIPQSSSLTLLCSLCFGDMENKLFARVQDQDGLLLRPVDDFLVTPHLDQ 872  
Db 820 IRGKSYVQCGIPQSSSLTLLCSLCFGDMENKLFARVQDQDGLLLRPVDDFLVTPHLDQ 879  
QY 873 AKTFLSTLVHGVPEYGCWMLNOKTWVNFVPEPGTGLGAAPYQLPALHCLFPWCGLLDTQT 932  
Db 880 AKTFLSTLVHGVPEYGCWMLNOKTWVNFVPEPGTGLGAAPYQLPALHCLFPWCGLLDTQT 939  
QY 933 LEVFCDSYGYAQTSTKTSITPQSVFKAGKTMWRNKLVSRLKCHGLFLDLQVNSLQTVCI 992  
Db 940 LEVQSDYSYARTSTRASLTFRNGFKAGRNRRKLFQVRLKCHSLFLDLQVNSLQTVCT 999  
QY 993 NIYKIFLLQAVRFHACVIOQLPFDORVRKNLTFFLGIISSOASCYAILKVKNFPMTKAS 1052  
Db 1000 NIYKIFLLQAVRFHACVIOQLPFDORVRKNLTFFLGIISSOASCYAILKVKNFPMTKAS 1059  
QY 1053 GS---FPPEAAHFWLCYQAFLLKLAHSAHYIKCLGLPLRTAQKLLCRKLPEATMTILKAAA 1109  
Db 1060 GAAGPLPSEAYQWLCHQAFLLKLTTRHRTVTVPLIGSLRTAQKLLCRKLPEATMTILKAAA 1119  
QY 1110 DPALSTDFQTILD 1122  
Db 1120 NPALPSPDFKTILD 1132  
RESULT 8  
US-10-208-243-2  
; Sequence 2, Application US/10208243  
; Publication No. US20030044394A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaeta, Federico C.A.  
; APPLICANT: Geron Corporation  
; TITLE OF INVENTION: Methods and Compositions for Eliciting an Immune  
; TITLE OF INVENTION: Response to a Telomerase Antigen  
; FILE REFERENCE: 015389-003500PC  
; CURRENT APPLICATION NUMBER: US/10/208,243  
; PRIOR FILING DATE: 2002-07-30  
; PRIOR APPLICATION NUMBER: US/09/675,321  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/112,006  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: WO PCT/US99/06898  
; PRIOR FILING DATE: 1999-03-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1132  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-208-243-2  
Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPRCRAVRSILRSRYREVWPLATFVRRLLGPEGRRRLVQCPDKIYRTLVAAQCLVCMHW 60  
Db 1 MPAPRCRAVRSILRSRYREVWPLATFVRRLLGPEGRRRLVQCPDKIYRTLVAAQCLVCPW 60  
QY 61 GSOPPPADLSFQVSSLSKELVARVVRQICERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
Db 61 DARPPAPAFSFRQVSKELVARVVRQICERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
QY 121 SYLNPNTVETLRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLYQICA 180  
Db 121 SYLNPNTVETLRVSGAMWLLSRVGGDILLVLLAHCAALYLLVPPSCAYQVCGSLYQICA 180

181 TTDIWPVSASVPRPVRNFTNLRFLQKSSROEAPKPLALPSRGTKXHLSTSTS 240  
181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
241 VPSAKKACVFPVREVEGP-----HQVLPTEPSGGKSW-VPSPARSPVPTAEKDLSSK 292  
232 LPLPKRPRGAPEPRTFVCGSWAHGRTGRSPDRGFCVVSAPAR-----PAEZATSL 286  
293 GKVSLSLS-GSVCKKXPSSTLSLPPRQNAFOLRP-FIETHFYSGDQGERLNPSF 350  
287 GALSCTRHSHPVSGHQHAGPSTSRPPRWDTPCPVYAEATKHFYSSGD-KEQLRPSF 345  
351 LLSNLQPNLTGABRLVETIFGSPRISGCLCRTHLSRBYWQWRPLFOQLLVNHAECQY 410  
346 LLSLRPSLTGABRLVETIFGSPRISGCLCRTHLSRBYWQWRPLFELLNHACQPY 405  
411 VLLSHCRFTANQVTDAL-----NTSPPHMLDLRLHSSPWQVY 452  
406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEEDTPRRLVQLLRHSSPWQVY 462  
453 GLPRLACLKVYSASLWGTNRHNERFFKNLKFISLGYKGLSLOELMWKMYEDCHLRS 512  
463 GVRACLRLVPPGLWGRHNERFLNRTKFKISLGHAKLSLOELTWKMSVRDCAWLR 522  
513 SPKDRVPAAEHRLERILATLFLWMDTYVQLLRFFVITESTFOKNLFFYRKSVM 572  
523 SPGVGVPAAEHRLERILATLFLWMDTYVQLLRFFVITESTFOKNLFFYRKSVM 582  
573 KLOSTGVQHLERVLRELSEVVRHODTWLAMPICLRFIPKNGLRPIVNMYSMGT 632  
583 KLOSTGIRQLKRVQLRELSAEVRQREARPAULTSLRFLPKPDGLRPIVNMVYVGA 642  
633 RALGRKQAOHTORLKTFLMNLVYRTKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
643 RTREREKAEALTSKXALFVLNVERARERGLGASVLGDDIHRARWTFVLVRAQDP 702  
693 TPRMTFVADVTGAYDAIPQKLVVEVANMTRHSESTYCIQVAVVRDQSGQVHKSF 752  
703 PPELVFKVVDVTGAYDTIPQDLTEVIAIIT-KPNTYCVRRYAVVQKAAHGHVRKAF 761  
753 QVTTLSLDLPVYQGLKHLQSDASALRNSVVIQSI SMNSESLSLDFLFLHLSVVK 812  
762 HVSTLTDLPVYQGLKHLQSDASALRNSVVIQSI SMNSESLSLDFLFLHLSVVK 819  
813 IGRCVYTCQGIPOGSSSTLSCSLCFGDMENKLF AEVQRDGLLRVDDFLVTPHL 872  
820 IRGKSVQCOGIPGSSSTLSCSLCFGDMENKLF AEVQRDGLLRVDDFLVTPHL 879  
873 AKTFLSTLVHGYPEYGCMLNLOKTVNPFVPGTILGGAAPVQLPAHCLFPWCGLLD 932  
880 AKTFLSTLVHGYPEYGCMLNLOKTVNPFVPGTILGGAAPVQLPAHCLFPWCGLLD 939  
933 LEVFCDSYGAOTSITSLTOSVFKAGTKMKNLLSVLRKCHKGLFLDLQVNSLTQVCI 992  
940 LEVQSDYSYARTSIRASITFNRGFKAGENMRKLFGLVRLKCHSLFLDLQVNSLTQVCT 999  
993 NIYKFLQAYFACVQVLPDQVRKLVLTFFLIGTISQASCCYATILKYNQWMLKAS 1052  
1000 NIYKILLQAYFACVQVLPDQVRKLVLTFFLIGTISQASCCYATILKYNQWMLKAS 1059  
1053 GS---PPPEAAHLCVQAPFLKLAHNSVYIKLGLPLRTAQKLCRKLPEATWTILKAAA 1109  
1060 GAAGPLPSEAVQWLCQAPFLKLTTRHYTVYVPLGLSLRTAQKLCRKLPEATWTILKAAA 1119  
1110 DPALSTDPOTILD 1122  
1120 NPALPSDFKILD 1132

## RESULT 9

US-10-054-295-225

Sequence 225, Application US/10054295

Publication No. US20030044953A1

GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Hatley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030044953A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,295  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
US-10-054-295-225

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 52.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPCPAVRSILRSRYREVWPLATFVRRLGPEGRLVQPGDKIYRTLVAQCILVCMHW 60  
Db 1 MPAPRCVRSILRSRYREVWPLATFVRRLGPEGRLVQPGDKIYRTLVAQCILVCMHW 60  
QY 61 GSQPPADLFSHQVSSLKELVARVORLCEHNRNVLAFGPELLNEARGGPPMATTSVR 120  
Db 61 DARPPAPSPRQVSCIKELVARVORLCEHNRNVLAFGPELLNEARGGPPMATTSVR 120  
QY 121 SYLNTVETLRVSGAWMLLSRVGDDLVLLAHALYLLVPPSCAYCVCGSPYQICA 180  
Db 121 SYLNTVETLRVSGAWMLLSRVGDDLVLLAHALYLLVPPSCAYCVCGSPYQICA 180  
QY 181 TTDIWPVSASVPRPVRNFTNLRFLQKSSROEAPKPLALPSRGTKXHLSTSTS 240  
Db 181 ATQARPPPHAS-GPRRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231  
QY 241 VPSAKKACVFPVREVEGP-----HQVLPTEPSGGKSW-VPSPARSPVPTAEKDLSSK 292



232 LPLPKPRGAAPEPRTFVGQSWAHQRTGSDRGFCVVSPAR-----PAAEATSL 286  
293 GKVSJLSL-GSVCKKPSSTLSPPRQNAQLRP-FIETRHFLXSGDQGRNLNPSF 350  
287 GALSCTRHSHPSVGRHAGPSTSPRPDPWTPCPVVAETKHFYSSGD-KEQLRPSF 345  
351 LLSNLQNTGARRLVEITFLGSRRTSPCLRTHLSRYWQMPFLFQOLLVNHACQY 410  
346 LUSLSRPSUTGARRLVETFLGSRPMPGTPRRLPRLPQRYWQMPFLFLELLGHAQCPY 405  
411 VRLRSHCRFTANQOQVTDAL-----NTSPHLMDDLRLHSSPWQVY 452  
406 GVLLKTHCPRAA--VTPAAGVCAKEKQGSVAAPBEDDPRVLQVLLQHSSPWQVY 462  
453 GFTBACLCKVVSASLWGRHRRRFFKXKIFISLKYKGLSLOELMWKQVDECHWLRS 512  
463 GFVRACLRRLVPPGLWGRHRRRFFKXKIFISLKYKGLSLOELMWKQVDECHWLRS 522  
513 SPKQDVPAAHRLRERILATFLWMDTVVQVLLRSFFYITESTFQKRLFFYRKSVWS 572  
523 SPQVGCVPAAHRLRERILATFLWMDTVVQVLLRSFFYITESTFQKRLFFYRKSVWS 582  
573 KLOSIGVROHLERVLRELQSEVRRHDTWAMPICRLRFPKNGRPIVNNYSNGT 632  
583 KLOSIGVROHLERVLRELQSEVRRHDTWAMPICRLRFPKNGRPIVNNYSNGT 642  
633 RALGRKQOQHTQRLKTLFSLMYERTKPHLMGSSVLGMDIYRTWRAFLVLRALDQ 692  
643 RTRRKRERLTSRKALFSLNVERARPPGLGASVLGDDIHRWRTFVLVRQDPP 702  
693 TPRLMYFKADVTGAYDAIPQGLVEYVANMIRHSESTYCIQYAVVRDSQGVHKSFR 752  
703 PPRLYFKVDVTGAYDTIPQDRLETVIASIK-PQNTYCVRYAVVQAAHGHVKAFKS 761  
753 QVTLSDIQPYMGOFKHLQSDASALRNSVVEQISVNSHSSSLPDPFLFLRHSVVK 812  
762 HVSTLTDLPQYMRQFVAHQET--SPLRDAVVEQSSSLNEASSGLFDVFLRFCHAVR 819  
813 IGRCYVTCQGGIPQGSSTLCSLCFCGDMENKLFPAVQRODGLLRFVDDLLVTPHLDQ 872  
820 IRKSVYVQCGIPQGSSTLCSLCFCGDMENKLFPAVQRODGLLRFVDDLLVTPHLDQ 879  
873 AKTFLTLVHGPEYGCMLNOKTVNFPVEPGTIGGAAPYQLPAHCLFPWCGLLLTQT 932  
880 AKTFLTLVHGPEYGCMLNOKTVNFPVEPGTIGGAAPYQLPAHCLFPWCGLLLTQT 939  
933 LEVPCDYSYVAOTSISLTSFQSVKAGTKMENKLLSVLRKCHGLFLDLQVNSIQTVCI 992  
940 LEVQSDYSYVAOTSISLTSFQSVKAGTKMENKLLSVLRKCHGLFLDLQVNSIQTVCI 999  
993 NIYKIFLLQAYRHACVIOQLPFDQVRKNTLFFLGIISQASCCYAILKVNPGMTLKAS 1052  
1000 NIYKIFLLQAYRHACVIOQLPFDQVRKNTLFFLGIISQASCCYAILKVNPGMTLKAS 1059  
1053 GS---PPERAAHLCYCAFLKLAHSHVYKCLIGPLTAQKLCRKLPEATMILKAAA 1109  
1060 GAAGPLPSEAVQWCHQAFLLKTRHRYVYVPLGSLRTAQQLSRKPLPGTLLTALEAAA 1119  
1110 DPALSTDFOTILD 1122  
1120 NPALPSDFKILD 1132

SULT 10

-10-054-611-225

Sequence 225, Application US/10054611

Publication No. US20030059787A1

GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.

Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. US20030059787A1el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,611  
FILING DATE: 18-Jan-2002  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/854,050  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1132 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
US-10-054-611-225  
Query Match 59.4%; Score 3505; DB 15; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;  
QY 1 MTRAPRCFPAVRSLLRSRYREVWPLATFVRRILGPEGRRLVQGPDKIYRTLVACLVCMHW 60  
DB 1 MPRAPRCRAVRSLLRSRYREVLPATFVRRILGPEGRRLVQGPDKIYRTLVACLVCMHW 60  
QY 61 GSQPPADLSFQVSSKELVARVVRQRLCERNERNVLAFGFELLNBARGGPPMAFTSSVR 120  
DB 61 DARPPAPAPSPRQVSCLELVARVQRLCERGAKNVLAFGFALLDARGGPPFAFTTSVR 120  
QY 121 SYLPNTVIETLRVSGAMWLLSRVGGDLLVLLAHCAVLLYLLVPSPCAVQVCGSPLVQICA 180  
DB 121 SYLPNTVTDALRGSGAWGLLRRVGGDLLVLLAHCAVLLYLLVPSPCAVQVCGSPLVQICA 180  
QY 181 TTDIWPVSASRYRTPRPGVGRNFTNLRFLQIKSSSRQEAAPKPLALPSRGTKRHLSTSTS 240  
DB 181 ATQARPPPPHAS-GPRRRLG-----CERAWNHSVREAGVPLGLPAGARRRGGSASRS 231  
QY 241 VPSAKKACYPVRVEEGP-----HRQVLPSPGKSW-VPSPARSPVEPTAEKILSSX 292  
DB 232 LPLPKRPRGAAPERTPVQGSWAHPGTRGSPDRGFCVVSAPAR-----PAAEATSL 286  
QY 293 GKVSJLSL-GSVCKKPSSTLSPPRQNAQLRP-FIETRHFLXSGDQGRNLNPSF 350  
DB 287 GALSCTRHSHPSVGRHAGPSTSPRPDPWTPCPVVAETKHFYSSGD-KEQLRPSF 345

```

351 LLSNLOPNTGARRVELIFLGSPPRTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 410
346 LLSLRPSLTGARRLVETIFLGSPPWPGTTRPLRPLPQRYWQMRPLFLELGNHAQCCY 405
411 VLLRSHCRFTANQQTVDAL-----NTSPHMLDLRLHSSPWQY 452
406 GVLLKTHCFLRAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRLVQLLRQSSPWQY 462
453 GFLRACLCVKVVSASLWGTNRHRRFFKNLKFISLGYKLSLOELMWKMKVEDDCHWLRS 512
463 GFVRACLRLVPPGLWGSRRHRRFLNTRKFKISLGHAKLSQELTWKVSVRDCAWLR 522
513 SPGKDRVPAAEHRLRERLTATFLWMDTVVQLLRFFYITESTFQKRLFFYKSVMS 572
523 SPGVCVPAAEHRLREELTAKFLHKLMSVYVVELLRFFYITETTFQKRLFFYKSVMS 582
573 KQSIGVROHLRVLRLRELSQEEVRRHQDTWLAMPICRLRFPKPNGLRPIVNMYSMGT 632
583 KQSIGIRQHLKRVQLRELSAEVRQREARFALLTSRLRFPKPDGLRPIVNMDDYVGA 642
633 RALGRKQAOHFTORLKTFLSMANVETKHPHMGSSVLGMNDIYRTWRAFLVRALDQ 692
643 RTFREKRAEELTSRKALFSVLYERARRPGLUGASVLGDDIHRARTFVLVRADQP 702
693 TPRMYFVKADVTGAYDAIPQGLVVEVVMNMRHSSTYCIQYAVVRRDSQGVHKSFR 752
703 PPELYFVKDVTGAYDTIPQDLRTEVIAIHK-PONTYCVRRYAVVQKAAGHVRKAFKS 761
753 QVTTLSLQPYMGOFLKHLQSDASALRNSVVEQSISNMNESSSLPDELHFLRHSVVK 812
762 HVSTLTLDQPYMRQFVAHQET--SPURDAVVEQSLSLNEASSGLDFVLRFWCHAVR 819
813 IGDRCYTQCGIPQGSSTLTLCSLCFQDMENKLFABEVORDGLLRFLVDDFLVTPHLQ 872
820 IRGKSVYQCGIPQGSSTLTLCSLCYDMENKLFAGIRRDGLLRVDDFLVTPHLTH 879
873 AKTFLSTLVHGVPEYGCMLNQLKTVNFPVPGTLGGAAPYQLPAHCLFWCCGLLDTQT 932
880 AKTFLRTLVRGPEYGCVMNLRKTVNFPVDEALGTAFAVMPAHGLFWCCGLLDTT 939
933 LEVFCDSYGAQTSIKTSITQSQVFKAGTWRNKLVLRLKCHGLFLDLQVNSLQTVCI 992
940 LEVQSDYSYARTSIRASLTNRGFKAGNRRKLFGLRLKCHSLFLDLQVNSLQTVCT 999
993 NTKYKFLQAVRHACVTLQSPDQVRNKLTFGLGITSQASCCYAILKVNPGMTLKAS 1052
1000 NTKYKILLQAVRHACVTLQSPHQVWKNPTFFLRVSDTASLCYSILKAKNAGMSLGAK 1059
1053 GS---FPPAAHLCVQAFLLKLAHSAHVYKCLLGLRLTAQKLCRLPEATWTILKAAA 1109
1060 GAAGLPSEAVQWUHQAFLLKLTFRVTVYVPLGLSLRTAQTLQSLKLPGLTTLTALAAA 1119
1110 DPALSTDFQTL 1122
1120 NPALPSDFKTL 1132

```

RESULT 11  
 S-10-105-963-2  
 Sequence 2, Application US/10105963  
 Publication NO. US2003006818A1  
 GENERAL INFORMATION:  
 APPLICANT: Geron Corporation  
 APPLICANT: Denning, Chris  
 APPLICANT: Clark, A. John  
 TITLE OF INVENTION: Animal Tissue with Carbohydrate Antigens Compatible for Human  
 TITLE OF INVENTION: Transplantation and a Carbohydrate Determinant Selection System  
 TITLE OF INVENTION: Recombination  
 FILE REFERENCE: 731/002  
 CURRENT APPLICATION NUMBER: US/10/105,963  
 CURRENT FILING DATE: 2002-03-21  
 PRIOR APPLICATION NUMBER: US 60/277,811

```

; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-105-963-2

Query Match      59.4%; Score 3505; DB 15; Length 1132;
Best Local Similarity 62.4%; Pred. No. 1.6e-316;
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCFAVRSLLRSRYREVWPLATFVERLQPEGRRLVQPDGPKIYRTLVACQICVCMHW 60
DB 1 MPAPRCFAVRSLLRSRYREVWPLATFVERLQPEGRRLVQPDGPAAPFALVAQCLVCVPM 60
QY 61 GSOPPPADLSFHQVSSKLKELVARVQVORLCERNERNVLAFGFELLNEARGGPPMATSSVR 120
DB 61 DARPPPAAPSFQVQSLKELVARVQLRCLERCAGKAVLAFGFALLDARGGPPFAFTTSVR 120
QY 121 SYLPTVETLRYSGAMWLLLSRVGDDILAVILLAHCALVLLVPPSCAYQVCGSPYQICA 180
DB 121 SYLPTVETDRLSGGAWGLLRVGVDDVLVHLLARCALFVLVAPSCAYQVCGPPYQIGA 180
QY 181 TTDTPSVSASRYPTPVGRNFTNLFQOIKSSRQEAQKPLALPSRGTKEHLSLTSTS 240
DB 181 ATQARPPPHAS-GPRLRLG-----CERAMNHSVREAGVPLGLPAPGARRRGGASRS 231
QY 241 VPSAKKACVYPRVEEGP-----HRQVLTPTSGKSW-VSPARSPEVPTAERDLASK 292
DB 232 LPUPKPRRGAAPPEPTPVQGSWAHPGRTGPDGFCVVSPPAR-----PAEATSLE 286
QY 293 GKVSDDLSS-GSVCCCHKPSSTLSLPPRQNAFQLRP-PIETRHFLYSRGDQGERLNPSF 350
DB 287 GALSGRHSHPSVGRQHAGPPSTSRPPRWDTPCPVYVAETKHFYSSGD-KEQLRPSF 345
QY 351 LLSNLOPNTGARRVELIFLGSPPRTSGPLCRTHRLSRRYWQMRPLFOQLLVNHAECQY 410
DB 346 LLSLRPSLTGARRLVETIFLGSPPWPGTTRPLRPLPQRYWQMRPLFLELGNHAQCCY 405
QY 411 VLLRSHCRFTANQQTVDAL-----NTSPHMLDLRLHSSPWQY 452
DB 406 GVLLKTHCFLRAA---VTPAAGVCAREKPGQSVAAPEEEDTDPRLVQLLRQSSPWQY 462
QY 453 GFLRACLCVKVVSASLWGTNRHRRFFKNLKFISLGYKLSLOELMWKMKVEDDCHWLRS 512
DB 463 GFVRACLRLVPPGLWGSRRHRRFLNTRKFKISLGHAKLSQELTWKVSVRDCAWLR 522
QY 513 SPGKDRVPAAEHRLRERLTATFLWMDTVVQLLRFFYITESTFQKRLFFYKSVMS 572
DB 523 SPGVCVPAAEHRLREELTAKFLHKLMSVYVVELLRFFYITETTFQKRLFFYKSVMS 582
QY 573 KQSIGVROHLRVLRLRELSQEEVRRHQDTWLAMPICRLRFPKPNGLRPIVNMYSMGT 632
DB 583 KQSIGIRQHLKRVQLRELSAEVRQREARFALLTSRLRFPKPDGLRPIVNMDDYVGA 642
QY 633 RALGRKQAOHFTORLKTFLSMANVETKHPHMGSSVLGMNDIYRTWRAFLVRALDQ 692
DB 643 RTFREKRAEELTSRKALFSVLYERARRPGLUGASVLGDDIHRARTFVLVRADQP 702
QY 693 TPRMYFVKADVTGAYDAIPQGLVVEVVMNMRHSSTYCIQYAVVRRDSQGVHKSFR 752
DB 703 PPELYFVKDVTGAYDTIPQDLRTEVIAIHK-PONTYCVRRYAVVQKAAGHVRKAFKS 761
QY 753 QVTTLSLQPYMGOFLKHLQSDASALRNSVVEQSISNMNESSSLPDELHFLRHSVVK 812
DB 762 HVSTLTLDQPYMRQFVAHQET--SPURDAVVEQSLSLNEASSGLDFVLRFWCHAVR 819
QY 813 IGDRCYTQCGIPQGSSTLTLCSLCFQDMENKLFABEVORDGLLRFLVDDFLVTPHLQ 872
DB 820 IRGKSVYQCGIPQGSSTLTLCSLCYDMENKLFAGIRRDGLLRVDDFLVTPHLTH 879

```

873 AKTFLSTLVHGVPEYGCINLQKTVNFFVPGTGLGGAAPYQIPAHCLFPWCGLLDITQT 932  
 880 AKTFLRTLVHGVPEYGCINLQKTVNFFVDEALGGTAFVQMPAHGLFPWCGLLDITRT 939  
 933 LEVFCDSYGAQTSIKTSITFQSVFVFRAGKTMMKLSVLRLKCHGLFLDLQVNSLQTVCI 992  
 940 LEVQSDYSSARTSIRASLTFFNKGFRAGNMKRLFGVLRKACHSLFLDLQVNSLQTVCT 999  
 993 NIYKIFLQAYRHACVIQPPDQVRKNTLTFPLGLISSQASCCVAILKVNKPMGNTLKAS 1052  
 1000 NIYKILLQAYRHACVLOLPPHQQVKNPTFFLRVISTASLCYSILKAKNAGMSLGAK 1059  
 1053 GS---FPPEAAHLCYQAFLLKLAHSLVYKCLLGLPRLTAQKLLCKLCKLPEATMTILKAA 1109  
 1060 GAAGLPSEAVQWLCHQAFLLKLTNRVTVVPLGSLRTAQQLSKLPGTITLTALEAA 1119  
 1110 DPALSTDFQITLD 1122  
 1120 NPALPSDFKITLD 1132

SULT 12

-10-044-692-2

Sequence 2, Application US/10044692

Publication No. US20030096344A1

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044,692

FILING DATE: 11-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1132 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-10-044-692-2

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCFAVRSLRSRVREYVPLATFVRRLGPGGRRLVQDGPDKIVRTILVAOCLVCMHW 60  
 DB 1 MTRAPRCFAVRSLRSRVREYVPLATFVRRLGPGGRRLVQDGPDKIVRTILVAOCLVCMHW 60  
 QY 61 GSCPPPADLSFHQVSSSLKELVARVVQRLCERNERNVLAQFELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPPPAAPSPRQVSKELVARVQLRCERGAKNVLAQFALLDARGGPPPEAFTTSSVR 120  
 QY 121 SYLPNTVLTETRVSGAWMLLSRVGDDLLVLLAHACALYLLVPSCAYOVCGSPLYOICA 180  
 DB 121 SYLPNTVTDALRGSGAWMLLSRVGDDVLLHLLARCALFVLVAQSCAYQVGGPLYOLGA 180  
 QY 181 TTDIMPSVSASYRTPRVPGRNFTNLRLFLQIKSSRQEPAPKPLALPSRGTKRHLSTSTS 240  
 DB 181 ATOARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSASRS 231  
 QY 241 VPSAKKARCYVPVRVEEGP-----HROVLPTSGKSW-VPSAPSPPEVPTAEKDLSK 292  
 DB 232 LPLPKRPRRGAPEPERTVQGGNAHPGRTRGSPFCVVSAPAR-----PABEATSLE 286  
 QY 293 GKVSDELISL-GSVCCCHKPSSSTLSLSPRQNAFQLRP-FIETRHFLYSRGDGOERLNPSF 350  
 DB 287 GALSCTRSHSPSVGQHAGPPSTSRPRPMDTFCPPVYAEKTHFLYSSGD-KEQLRPSF 345  
 QY 351 LLNSLQNLTGARLVEIIFLGSRRSGPLCRTHRLSRRYQWRPLFQOLLNVHAEQY 410  
 DB 346 LLSSLRPSLTGARLVEIIFLGSRRSGPLCRTHRLSRRYQWRPLFQOLLNVHAEQY 405  
 QY 411 VLLRSHCRFTANQQVTDAL-----NTSPHLLMDLLRLHSSPQVY 452  
 DB 406 GVLLKTHCPRLAA---VTPAAGVCAREKPGQSVAPBEEDTDPRLLVQLLRQHSPPQVY 462  
 QY 453 GFLRACLCKVWSASLWGRHNERFFKNLKFISLGYKGLSLQELMMKMKVBDCHWLRS 512  
 DB 463 GFVRACLRLVPPGLWGRSHNERFLNKKFISLGHAKLSLQELTWKMSVRDCAWLR 522  
 QY 513 SPKDRVPAAEHRLERILATFLWMDTVYVQLRSPFYITESTFQKRLFFYRKSWS 572  
 DB 523 SPVGCVPAAEHRLERILAKFLHWSVYVVELLRSFFYTETTFQKRLFFYRKSWS 582  
 QY 573 KLSIGVRQHLRVRRLRELSQSEVVRHHQDTWLAMPICRLRFIPKPNGLRPIVNMYSMGT 632  
 DB 583 KLSIGIRQLKRVQLRELSAEVVRQREARPAALLTSRLRFIPKPDGLRPIVNMYYVGA 642  
 QY 633 RALGRKQAOHFTQRLKTLPSMLNRYERTKEPHLMSVGLGNDIYRTWRAPVLRALDQ 692  
 DB 643 RTRREKAEALTSRVKALFSLVNYERARRPGLLGASVGLGDDIHRARWTFVLRVRAQDP 702  
 QY 693 TPRMYFVRADVTGAYDAIPQOKLVEVVAVMIRHSESTYICIRQAVVRRDSQGVHKSFR 752  
 DB 703 PPELYFVKVDVTGAYDTIPQDRLTEVIAIIK-PQNTYCVRRYAVVQKAAHGVRAKFS 761  
 QY 753 QVTTLSDIQPYMGQFLKHLQSDASALRNSVVIQSSISMNESSSLDFPFLHFLRHVVK 812  
 DB 762 HVTSLTDLQPYMRQFVAHQET--SPLRDVAVIEQSSSLNEASSGLDFVLRPMCHAVR 819  
 QY 813 IGDRCVTCQGIPOGSSSLTLLCSLCFCGDMENKLFASFVORDGLLLRFVDDFLVTPHLDQ 872  
 DB 820 IRGKSYVQCGIPIQSSILSTLLCSLCGDMENKLFAGIRDDGLLLRLVDDFLVTPHLDQ 879  
 QY 873 AKTFI-STLVHGVPEYGCINLQKTVNFFVPGTGLGGAAPYQIPAHCLFPWCGLLDITQT 932

```

b 880 AKTFRTLVRGVPEYGCNVLRKTVNPFVDEALGCTAFVQMPAHGLFPWCGLLDTRT 939
Y 933 LEVFCDSYSGYACTSIKTSILTSQSVFKAGKTRNKLLSVLRKCHGLFLDLQVNSLOTVC 992
b 940 LEVQSDYSYARTSIRASLTNRGFKAGRNWRKLFGLRLKCHSLFLLDLQVNSLOTVC 999
Y 993 NYIKFLQAVRHACVITQLPDDQVRNKLTFGLIGISSQASCCYAILKYNPMTLKAS 1052
b 1000 NYIKILLQAVRHACVITQLPDDQVRNKLTFGLIGISSQASCCYAILKYNPMTLKAS 1059
Y 1053 GS---PPEAAHLCYQAFLLKLAHSHVYKCLGLPLRTAQKLLCRKLPEATWTILKAAA 1109
b 1060 GAAGLPBPAVOWLCHQAFLLKLRHRTVYVPLLSLRTAQQLSRKLPGLTTLTALAAA 1119
Y 1110 DPALSTDFQTILD 1122
b 1120 NPALPSDFKTILD 1132

```

## RESULT 13

S-10-044-539-2

Sequence 2, Application US/10044539

Publication No. US20030100093A1

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/10/044,539

FILING DATE: 11-Jan-2002

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

```

; LENGTH: 1132 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-044-539-2

```

Query Match 59.4%; Score 3505; DB 15; Length 1132;  
 Best Local Similarity 62.4%; Pred. No. 1.6e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

```

QY 1 MTRAPRCPAVRSILRLSRKYREWPLATFVRRLQEGRRLLVQPGDPIKYRLVAQCLVCMHW 60
DB 1 MPRAPRCAVRSILRLSRKYREWPLATFVRRLQEGRRLLVQPGDPIKYRLVAQCLVCMHW 60
QY 61 GSQPPPADLRFHQVSSIKELVARVQRLCERNERNVLAQFELLNEARGGPPMATSSVR 120
DB 61 DARPPPAAPSRQVSCIKELVARVQRLCERNERNVLAQFELLNEARGGPPMATSSVR 120
QY 121 SYLPTVTIETLRVSGAWMLLSRVGDDLLVYLLAHALVLLVPPSCAYQVCGSPLYQICA 180
DB 121 SYLPTVTIETLRVSGAWMLLSRVGDDLLVYLLAHALVLLVPPSCAYQVCGSPLYQICA 180
QY 181 TTDIWPVSASVYRTPPVGRNFTNLRFLOQIKSSSRQEAQKPLALPSRGTKXHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRRGGSASRS 231
QY 241 VPSAKKARCYVPRVREGP-----HRQVLPSPGKSW-VSPARSPEVPTAEKDLSSK 292
DB 232 LPLPKRPRGAPEPERTPVQGSWAHPGRTGRGFCVVSFAR-----PAEATSLLE 286
QY 293 GKVSDDLSS-GSVCCCKHKPSSTLSLSPPRQNAQLRP-FIETRHFLYSGDQOERINPSF 350
DB 287 GALSGRHSHPVGRQHAGPPSTSRPPRWDTPCPVVAETKHFYSSGD-KEQLRPSF 345
QY 351 LLSNLQNLTGARLVEIIFLGRPRTSGPLRTHLSRRYQWQMRPLFOQLLVNABCOY 410
DB 346 LLSLSRPSLTGARLVETIIFLGRPRTSGPLRTHLSRRYQWQMRPLFOQLLVNABCOY 405
QY 411 VLLRSHCRFTANQOVTDAL-----NTSPPHMLDLRLHSSFPWOY 452
DB 406 GVLLKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLVQLLRQHSFPWOY 462
QY 453 GLFRACLCKVVSASLWGTNRNERRFFKNLKKTSISLKYKLSLOELMHWKVEDCHWLRS 512
DB 463 GFVRACLRLRVLPPGLWGRNERRFLNKKTSISLKYKLSLOELMHWKVEDCHWLRS 522
QY 513 SPGKDRVPAAEHLRREILATFLFMDTVVYVOLLRSFYITESTFOKNRFFYKSVWS 572
DB 523 SPGVGCVPAAEHLRREILAKFLHWSVTVVELGRSFYVTTTFOKNRFFYKSVWS 582
QY 573 KLSIGVROHLRVRRLRELSQEEVRHHQDTLWAMPTICRLRFIPKPNGLRPIVNMYSMT 632
DB 583 KLSIGVROHLRVRRLRELSQEEVRHHQDTLWAMPTICRLRFIPKPNGLRPIVNMYSMT 642
QY 633 RALGRKQAOHQFTORLKTFLSMLNYERTKPHLMSSVLGMNDIYRTWEAFVLRVRLDQ 692
DB 643 RTFRKRAERLTSRVKALFVNLNERRRPPGLGASVLGDDIHRAMWTFVLRVRAQDP 702
QY 693 TPRMYFVKADVTGAYDAIPQGLVEVAVNMIRHSESTYCIROYAVVRDRSQOVHKSFR 752
DB 703 PPELYFVKADVTGAYDTIPQDLTEVIAIHK-PQNTYCVRRYAVVQKAAGHGVKAFKS 761
QY 753 QVTLSDLOPYMQOFLKHLQSDASALRNSVITEQISNMNNESSSLFDFFLHRSVVK 812
DB 762 HVSTLTDLPYMRQFVAHQET--SPLRDVAVTEQSSLSNEASSGLFDVFLRFMCHAVR 819
QY 813 IGDRCYTQCQGIPOGSSLSLTLCSLCFGDMENKLPABVQVDGLLRFDVDFLLVPHLDQ 872
DB 820 IRGKSYVQCQGIPOGSSLSLTLCSLCFGDMENKLPABVQVDGLLRFDVDFLLVPHLDQ 879
QY 873 AKTELSTLVHGVPEYGCNVLRKTVNPFVDEALGCTAFVQMPAHGLFPWCGLLDTRT 932

```

880 AKTFLRLTVRGVPEYCVNLKRTVVNFFVEDEALGGTAFTVOMPAHGLFPWCGLLDTRT 939  
 933 LEVFCYSGVQTSIKTSITFOSVFKAGTKMTNKLSSLVRLKCHGLFLDLQVNSLQTVCI 992  
 940 LEVOSDYSSYARTSIRASLTFRNGFKAGRMERKLFGLVRLKCHSLFLDLQVNSLQTVCT 999  
 993 NIYKIFLQAYREHACVIOLEPDORVRKNLTFFLGIISQASCCVAILKVNKPGWTLKAS 1052  
 1000 NIYKILLQAYREHACVIOLEPHQVQWKRKPTFFLRVISTASLCSYILKAKNAGNSLGAK 1059  
 1053 GS---PPPEAAHWLCYQAFLLKLAHSVYKCLLGLPLRTAQKLLCKRLPEATMILKAAA 1109  
 1060 GAAGLPSEAVOMLCHQAFLLKLTREHVVYVPLGSLRRTAQQLSKLPGTTLTALFAAA 1119  
 1110 DPALSTDFQTILD 1122  
 1120 NPALPSDFKTILD 1132

## SULT 14

-10-044-692-323

Sequence 323, Application US/10044692

Publication No. US20030096344A1

## GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044,692

FILING DATE: 11-Jan-2002

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912,951

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US 08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851,843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 323:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 amino acids

TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 323:  
 US-10-044-692-323

Query Match 59.4%; Score 3505; DB 15; Length 1154;  
 Best Local Similarity 62.4%; Pred. No. 1.7e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCAPVARSILRSRYREVMPLATFVRRLLGPEGRRLVQDPDKIYRTLVVAQCLVCMHW 60  
 DB 1 MPRAPRCVARSLLASHREVLPATFVRRLLGPOGWRVLRQGDPAAFALVAQCLVCPW 60

QY 61 GSQPPADLSFHQVSSLSKELVARVVRQRLCERNERNVLAFGFELLNEARGGPPNFTSSVR 120  
 DB 61 DARPPPAAPSPRQVSCLEKELVARVLRQRLCERGAANVLAFGFALLDGGARGGPPPEAFTTSSVR 120

QY 121 SYLPNTVTETLRVSGAWMLLSRVGDDLLVLLAHACALYLLVPSCAYQVCGSPLYOICA 180  
 DB 121 SYLPNTVTDLRGSGAWGLLRVGVDDVLLHLLARCALFVLVAFSCAYQVCGSPLYOLGA 180

QY 181 TTDIWPSVSAHYRTPVPGRNFTNLRFLQIKSSSRQEPAPKPLALPSRGTGRHLSTSTS 240  
 DB 181 ATCARPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSSASRS 231

QY 241 VPSAKKARCYPVPRVEGPG-----HRQVLPSTGKSW-VPSPARSPVEVPTAEKDLSSK 292  
 DB 232 LPLPKRPRGGAPEPERTVGGGSHAGPGRTRGSDRGFCVVSAPAR-----PAEATSL 286

QY 293 GKVSLLSLSS-GSVCCCHKPSSSTLSLSPRQNAFQRLP-FIETRHFLYSRGGDGERLNPSF 350  
 DB 287 GALSCTRHSHPSGVGRQHHAGPPSTSRPRPMDTPCPVYAETKHFLYSSGD-KEQLRPSF 345

QY 351 LLSNLOPNLTGARRLVEIIFLGSRRPSGCLRTHLSRRYQWNRPLFQOLLVNHACQY 410  
 DB 346 LLSLRPLSTGARRLVETIFLGSRRPMMGTTPRRLPRLPQRYWQMRPLFLELLGNHACQY 405

QY 411 VLLRASHCRFRATANQVTDAL-----NTSPHLLMDLLRLHSSSPQVY 452  
 DB 406 GVLLKTHCFELRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLLVQLLRQHSPPQVY 462

QY 453 GFLRACLCKVWSASLWGRTHNRERRFFKNLKEFIISLGKYKLSLOELMWKMKVEDCHWLRS 512  
 DB 463 GFVRACLRLRVLPPGLWGRSHNRERRPLRNTKKFISLGKXAKLSLOELTWQMSVRDCAWLR 522

QY 513 SPKDRVPAAEHRLERILATFLWMDTYVQLLRGFEYITESTFQKNRLFVYRKSVMS 572  
 DB 523 SPGVGCVPAAEHRLERILAKFLHLMVSVYVELLRSFYVYTTTFQKNRLFVYRKSVMS 582

QY 573 KLQSIGVROHLERVRRLRELSQBEVRHHQDTWLAMPICRLRFIPKPNGLRPTVNNMSYSGT 632  
 DB 583 KLQSIGIRQHLKRVQLRELSEAEVQREARPAALLTSRLRFIPKPDGLRPIVNMVYVGA 642

QY 633 RALGRKQAHQHTQRLKTLFSLNRYERTKPHLMGSSVLGMNDIYRTWRAFLVLRALDQ 692  
 DB 643 RTFREKKAERLTSRVKALFSLNRYERARRPCLLGASVLGDDIHRAMRTFLVLRVRAQDP 702

QY 693 TRPMYFVKADVTGAYDAIPOGKLVEVVMIRHSESTYCIQYAVAVRDSQGVKSKFRR 752  
 DB 703 PELVFKVDVTGAYDTIPQDLTEVIAIILK-PQNYICVRYAVVQKAAHGHVYKAFKS 761

QY 753 QVTTLSDLQPYMGQFLKHLQSDASALRNSVVIQSSISMNESSSLDFFLHFLHRSVVK 812  
 DB 762 HVSSTLDLQPYMRQFVAHLQET--SPLADAVVIEQS8SLNEASSGLEFDVFLRFMCHAVR 819

QY 813 IGDRCVTCOGIPQGSLSLTLCSLCFGDMENKLFQAEVORDGILLRVDVDFLLVTPHLQ 872  
 DB 820 IKGSTVQCQGIPOGSIILSTLLCSLCYGD MENKLFAGIRDRGLRLVDDFLVTPHLTH 879

QY 873 AKTFIISTLVHGVPEYGCMLNLOKTYVNFVFPFGTLGGAAAPYQLPAHCLFPWCGLLDTQT 932

880 AKTFLRLTVRGVPEYGCVNLRKTVNFPVEDEALGGTAFVQMPAHGLFPWCGLLDTRT 939  
 933 LEVFCYSGVAQTSIKTSTFQSFVAGKTMKNKLLSVLRKCHGLFGLDQVNSLQTVCI 992  
 940 LEVQSDYSSVARTSIRASLTFRNGFRAGRMRRKLFVGLRKLKCHSLFGLDQVNSLQTVCT 999  
 993 NIYKIFLQAYRHACVIGLPDQVRKNTLTFPLGIISSQASCCVAILKVKPNQNTLKAS 1052  
 1000 NIYKILLQAYRHACVIGLPDQVRKNTLTFPLGIISSQASCCVAILKVKPNQNTLKAS 1059  
 1053 GS---FPPEAAHLCYQAFLLKLAHSHVIYKCLLGLPLRTAQKLLCKLPEATMTILKAAA 1109  
 1060 GAAGPLSEAVQMLCHQAFLLKLTREHVIYVPLGLSLRTAQQLSEKLPGLTTLTALEAAA 1119  
 1110 DPALSTDFQILD 1122  
 1120 NPALPSDFKILD 1132

# 38ULT 15

3-10-044-539-323  
 Sequence 323, Application US/10044539  
 Publication No. US20030100093A1  
 GENERAL INFORMATION:

APPLICANT: Cech, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morlin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND

THERAPEUTIC METHODS

NUMBER OF SEQUENCES: 335

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/044.539

FILING DATE: 11-Jan-2002

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/912.951

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/854.050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/851.843

FILING DATE: 06-MAY-1997

APPLICATION NUMBER: US 08/846.017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844.419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724.643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002600US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 323:

SEQUENCE CHARACTERISTICS:

LENGTH: 1154 amino acids

TYPE: amino acid  
 STRANDEDNESS: <Unknown>  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 323:  
 US-10-044-539-323

Query Match 59.4%; Score 3505; DB 15; Length 1154;  
 Best Local Similarity 62.4%; Pred. No. 1.7e-316;  
 Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCAPAVRSLRSRYREVMPLATFVRRLLGPEGRLLVQDPKTYRTLVLAQCLVCMHW 60  
 DB 1 MTRAPRCAPAVRSLRSRYREVMPLATFVRRLLGPEGRLLVQDPKTYRTLVLAQCLVCMHW 60

QY 61 GSQPPPADLSFQVSSLSKELVARVQRLCERNERNVLAFGPELLNEARGGPPMAFTSSVR 120  
 DB 61 DARPPPAAPSFQVSCLEKELVARVQRLCERGAQVLAFGFALLDARGGPPPEAFTTSVR 120

QY 121 SYLPNTVIELTRVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPYQICA 180  
 DB 121 SYLPNTVIELTRVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPYQICA 180

QY 181 TTDIWPVSASVYRTPRPGVGRNFTNLRFLQOKSSSRQBPAPKPLALPSGTYKXHLSTST 240  
 DB 181 ATQAREPPPHAS-GPRRLG-----CERAWNHSVREAGVPLGLPAPGARRRGGSASRS 231

QY 241 VPSAKKARCYPPRYVEEGP-----HQVLPTPSGKSW-VPSPARSPVEVTAEKDLSK 292  
 DB 232 LPLPKPRRGAAPERFIPVGGQSWAHFGRTRGSDRGFCVVSAR-----PAEATSLE 286

QY 293 GKVSIDLSL-GSVCCCKHPSSTLSLSPRONAFQLRP-FIETRIPLYSRGGQRLNPSF 350  
 DB 287 GALSGRTHSHPSVGRQHAGPPSTSRPPMDTPCPVYAEKHFLLYSSGD-KEQLRPSF 345

QY 351 LLSNLQNLTGARLLVEIIFLGSRRPTSGPLCRHLRSRYWQMEPLFOQLLVNHAECQY 410  
 DB 346 LLSLSRPSLTGARRLVETIFLGSRRPMEGTFRRLPRLPQRTQWRPLLELLGHAQCPY 405

QY 411 VRLLRSHCRFTANQQVTDAL-----NTSPPHLMDLLHSHSPWQVY 452  
 DB 406 GVLLKTHCPLRAA---VTPAAGVCAREKPGQSVAAPEBEDTPRRLVQLLRQHSSPWQVY 462

QY 453 GFTRACLCVKVVSASLWGRHNRERFFKNLKFISLGKYKLSLOELWMKMKVEDCHMLRS 512  
 DB 463 GFVTRACLRLLVPEGLWGRHNRERFLNFKFISLGKHAKLSQELTWNKSVRDCAMLR 522

QY 513 SPQKDRVPAAEHRLRERILATLFWLMDTVVQVLLRSPFYITESTFQKNRFFYRKSVMS 572  
 DB 523 SPGVGCVPAAEHRLRERILATLFWLMDTVVQVLLRSPFYITESTFQKNRFFYRKSVMS 582

QY 573 KLOSIGVROHLEVRVRELSEVHRHQDTWLAMPICRLRIPKPNGLRPIVNNYSNGT 632  
 DB 583 KLOSIGTRQHLKRVQLRELSEAEVRQREARPAALLTSRLRFPKPDGLRPIVNNMDYVGA 642

QY 633 RALGRKQAOCHFTQRLKTLFSLMNVERTKPHLMGSSVLGMNDIYRTWRAFVLRVRLDQ 692  
 DB 643 RTRERKRAELTSRVKALFVSLNVERARPGLLGASVLGLDDIHRANRTFVLRVRAQDP 702

QY 693 TPRMYFYKADVTGAYDAIPQGLVEVAMNHRSSSTYCIROYAVVRDSDGQVHKSPRR 752  
 DB 703 PPFLYFVKVDVTGAYDTIPQDRLTEVIAIIR-PONTYCVRYAVVQAAHGHVRAKFS 761

QY 753 QVTTLSLPQYMGQFLKHLQDSDASALNSVVIEGSIWNSSSSLSLFDFFLHFLRHSVVK 812  
 DB 762 HVSFLTDLQYMEQFVAHQEI--SPLRDAVIEGSSLSNEASSGLFDVFLFMCHAVR 819

QY 813 IGRCYTQCQGIPOGSSLSLTLCSLCFGDMENKLFABVQDGLLRLFVDDFLLVTPHLDQ 872  
 DB 820 IGRKSYVQCQGIPOGSSLSLTLCSLCYGDENKLFAGIRRDGLLRLVDDFLLVTPHLDQ 879

QY 873 AKTFELSLVHGVPYEGCMINLOKTVNFEVPEGTGGAPYQLPAHCLFPWCGLLDQ 932

```

880 AKTFLRLVRGPEYCCVNLKRTVVNPFVEDEALGGTAFVQMPAHGLFPWCGLLLDTRT 939
933 LEVFCDYSGYAQTSIKTSLTFSVPKAGKTMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
940 LEVQSDYSYARTSIRASLTENRPFKAGNMRRKLFGLVLRKCHSLFLDLQVNSLQTVCT 999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
993 NIYKIFLLQAYRHCACVIOLEPQORVRKULTFELGISSQASCCYAILKVNPGMTLKAS 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1000 NIYKILLQAYRHCACVLOLPHQVQVKNRPTFLRVISDTASLCYSILKAKNAGMSLGAK 1059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1053 GS---FPEEAAHWLCYQAFLLKLAHSHVIYKCLLGLPLRTAQKLLCRKLPEATMTILKAAA 1109
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1060 GAAGPLPSEAVQWLCYQAFLLKLTNRVTVYVLLGSLRTAQTLRSRKLPGTTLTALAAAA 1119
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1110 DPALSTDFQTILD 1122
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1120 NPALPSDFXTILD 1132
    : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

arch completed: February 4, 2004, 14:17:56  
b time : 50 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

n on: February 4, 2004, 14:09:14 ; Search time 27 Seconds  
(without alignments)  
3996.343 Million cell updates/sec

tie: US-09-042-460-2  
effect score: 5901  
quence: 1 MTRAPRCPAVRSLRSRYE.....TLKAAADPALSTDFQILD 1122

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

arched: 283308 seqs, 96168682 residues

tal number of hits satisfying chosen parameters: 283308

nimum DB seq length: 0  
ximum DB seq length: 2000000000

st-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : PIR\_76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result	No.	Score	Query Match	Length	DB	ID	Description
1	3505	59.4	1132	2	T03844	telomerase catalyt	
2	777.5	13.2	1123	2	T51517	telomerase reverse	
3	624	10.6	989	2	T03838	telomerase catalyt	
4	426.5	7.2	1132	2	T31107	telomerase reverse	
5	401.5	6.8	1117	2	T14891	telomerase (EC 2.7	
6	341	5.8	884	2	S53396	telomerase catalyt	
7	133.5	2.3	561	2	T20392	hypothetical prote	
8	131.5	2.2	330	2	E98119	transposase, uncha	
9	131	2.2	778	2	I38487	tastin - human	
10	129	2.2	502	2	T24227	hypothetical prote	
11	125	2.1	1529	2	T02730	RNA-directed DNA p	
12	124	2.1	4540	2	T30838	cytoplasmic dynein	
13	121	2.1	2088	2	E71436	hypothetical prote	
14	120.5	2.0	1646	2	T20740	hypothetical prote	
15	116.5	2.0	544	2	T11216	reverse transcript	
16	116	2.0	881	2	E82097	protein-P-II uridy	
17	115.5	2.0	2157	2	S71461	proline-rich prote	
18	114.5	1.9	1587	2	AB2012	hypothetical prote	
19	111.5	1.9	1415	1	EDBEGA	immediate-early pr	
20	111.5	1.9	2606	2	T03159	large tegument pro	
21	110	1.9	907	2	A24938	hypothetical T2 pr	
22	109.5	1.9	710	2	T41586	hypothetical serin	
23	109.5	1.9	1214	2	T00356	hypothetical prote	
24	109.5	1.9	2493	2	T40540	hypothetical prote	
25	109.5	1.9	3147	2	T21328	hypothetical prote	
26	108.5	1.8	351	2	S50754	hypothetical prote	
27	108.5	1.8	1147	2	T42627	ADP-ribosylation f	
28	108	1.8	880	2	S49627	regulatory protein	
29	108	1.8	1474	2	B85188	retrotransposon li	

30 107.5 1.8 600 2 A52570  
31 107 1.8 510 2 H84824  
32 107 1.8 949 2 T38543  
33 107 1.8 1003 2 E85432  
34 107 1.8 1121 2 A82809  
35 107 1.8 1456 2 T01397  
36 106.5 1.8 1256 2 C71436  
37 106 1.8 584 2 S40013  
38 106 1.8 877 2 T43449  
39 106 1.8 1017 2 T48452  
40 106 1.8 1115 2 A45761  
41 106 1.8 2528 2 T20719  
42 105.5 1.8 745 2 H85048  
43 105.5 1.8 811 2 T25166  
44 105.5 1.8 3998 2 S57437  
45 105 1.8 1493 2 T10757

## ALIGNMENTS

## RESULT 1

T03844  
telomerase catalytic chain - human  
N;Alternate names: telomerase reverse transcriptase  
C;Species: Homo sapiens (man)  
C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 08-Oct-1999  
C;Accession: T03844  
R;Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner, S.  
Science 277, 955-959, 1997  
A;Title: Telomerase catalytic subunit homologs from fission yeast and human.  
A;Reference number: Z15111; MUID:97400623; PMID:9252327  
A;Accession: T03844  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1132 <NAK>  
A;Cross-references: EMBL:AF015950; NID:g23330016; PIDN:AACS1672.1; PID:g2330017  
A;Experimental source: kidney  
C;Genetics:  
A;Gene: TRT  
A;Map position: 5p

Query Match 59.4%; Score 3505; DB 2; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 1.9e-252;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

QY 1 MTRAPRCPAVRSLRSRYEYVPLATFVRRILGPGRLVQDPKIVRTLVACLVCMHW 60  
Db 1 MTRAPRCPAVRSLRSRYEYVPLATFVRRILGPGRLVQDPKIVRTLVACLVCMHW 60

QY 61 GSQPPADLSFHQVSSSLKELVARVQRLCERNERNVLAFFELLINEARGGPPMAFTSSVR 120  
Db 61 DARPPPAAPSPQVSCLEKELVARVQRLCERNERNVLAFFELLINEARGGPPMAFTSSVR 120

QY 121 SYLPTVITETRVSGAWMLLSRVGDDLLVLLAHCALYLLVPSCAYVCGSPLYQICA 180  
Db 121 SYLPTVITDALRGSGAWMLLSRVGDDLLVLLAHCALYLLVPSCAYVCGSPLYQICA 180

QY 181 TTDIWPSVSASYPTRPVGRNFTNLRILQIKSSSRQEAAPKPLALPSRGTKRHLSLTSTS 240  
Db 181 ATCAARPPPHAS -GPERRLG-----CERAWNHVSREAGVFLGLPAGARRRGSSASRS 231

QY 241 VPSAKARCPYPRVTEGP-----HQVLPTSGKSW-VPSARSPPEVTAEDLSSK 292  
Db 232 LFLPKRPRGAPPEPTEPFGQSWAHFGRTRGSDRGFCVSPAR-----PABEATSL 286

QY 293 GKVSDELISL-GSVCKHKPSTSLSPQNAFQLRP-FIETRHLYSRGDGGRINPSP 350  
Db 287 GALSSTRSHSPSGVGHAGPSTSRPRPMDTCCPPVYAEKHFLLYSSGD-KEQLPSP 345

QY 351 LLSNTQPNLTGARLVEIFIGSRPRTSGPLCRTHRLSRYYQWRPFLQOLLVNHAEQY 410  
Db 346 LLSLSRLPSTGARLVEIFIGSRPRTSGPLCRTHRLSRYYQWRPFLQOLLVNHAEQY 405



Y 411 VLLRSHCHFRFANQOVTAL-----NTSPHLLMDLLRLHSSPWQVY 452  
b 406 GVLLKTHCPRAA---VTPAAGVCAREKPGQSVAAPEEDDPRLVOLLRQHSPPWQVY 462  
Y 453 GFLRACLCKVVSASLWGTNRHNERFFKNLKFISLGKVGKLSLOELMMKMKVDDCHWLRS 512  
b 463 GVRACLRRLVPPGWSGRHNERFLNKKFISLGKHAHLSLOELTWKMSVRDCAWLRR 522  
Y 513 SPGRDVPAAERHLRERILATFLFWMDTYVYVOLLRSFFYITESTFOKNLFFYRKSVM 572  
b 523 SPGVGCVPAERHLREELILAKFLHLMVSVYVVELLRSFFYVTTTFQKNLFFYRKSVM 582  
Y 573 KLOSTGVORHLERVLRLBELSQEVRHODTWLAMPICELREIPKPNGLRIVNMSVSGT 632  
b 583 KLOSTGIROHLKRVQLRELSEAEVRQREARPALTLKSLRPIPKPDGLRPIVNDYVYGA 642  
Y 633 RALGRKQOQHTQRLKTLFSLMNTYRTKPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692  
b 643 RYFREKRAERLTSVKALFSLVNYERARFGLLGASVLGDDIHRARWTFVLVRAQDP 702  
Y 693 TRPMYFKADVTGAYDAIPQGLKLEVVANMIRHSESTYCIROYAVVRDSSQGVHKSFR 752  
b 703 PPELYPKVDVTGAYDTIPQDRLTVEIASIHK-PQNTYCVARYAVWQAAHGHVRKAFKS 761  
Y 753 QVTLSDLOPYMGQFLKLODSASALRNSVYVIEOSISMNESSSLDFFLHFLRHSHVK 812  
b 762 HVSITLTDLPYMQFVAHLOET--SPLRDVAVIEOSSSLNEASSGLFDVFLRFCHAVR 819  
Y 813 IGRDVCYQCGIPQSSSLTLCSCFCGDMENKLPFAVORGLLLRFDVDFLLVTPHLQ 872  
b 820 IRKSVYVQCGIPQSSSLTLCSCYGMENKLPAGIRRDGLLRVLDVDFLLVTPHLTH 879  
Y 873 AKTFSLTVHGVPEYCGMNLCKTVNPPVEPTLGGAAPOLPALCLFPMCGLLDTOT 932  
b 880 AKTFSLTVHGVPEYCGVNVNLRKTVNPPVEDEALGTAFOVMPAHLFPWCGLLDTRI 939  
Y 933 LEVFCDSYGAQTSIKTSLTQSFVKAGTKMENKLLSVLRKCHGLFLDLQVNSLQTVCI 992  
b 940 LEVQSDYSVARTSIRASLTFRNFKAGNRKLPFGVLRKCHSLFLDLQVNSLQTVCT 999  
Y 993 NIYKIFLQYRFHACVQLPQDORVKLTFPLGIISSQASCCVAILKVNQPMTLKAS 1052  
b 1000 NIYKILLQYRFHACVQLPQHQVQWKNPTFFLRVSDTASLCYSILKAKNAGVSLGAK 1059  
Y 1053 GS---PFPAAHMLCYQALLKLAHSVYKCLLQPLRTAQKLLCRKLPEATMTILKAA 1109  
b 1060 GAAGPILPSEAVQNLCHQAFLLKLRHVTVYVPLGSLRTAQTLRSKLPGLTTLALEAA 1119  
Y 1110 DPALSTDFQTLTD 1122  
b 1120 NPALPSDFKTLTD 1132

RESULT 2  
T51517  
c:elomerase reverse transcriptase - Arabidopsis thaliana  
v:Alternate names: protein F5E19\_190  
s:Species: Arabidopsis thaliana (mouse-ear cress)  
d:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
a:Accession: T51517  
s:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Men  
submitted to the Protein Sequence Database, August 2000  
a:Reference number: Z25394  
a:Accession: T51517  
a>Status: preliminary  
a:Molecule type: DNA  
a:Residues: 1-1123 <SAT>  
a:Cross-references: EMBL:AL391147  
a:Experimental source: Cultivar Columbia; BAC clone F5E19  
c:Genetics:  
a:Map position: 5  
a:Introns: 100/3; 125/3; 147/3; 185/1; 300/3; 325/1; 369/2; 414/3; 765/3; 942/2; 1033/2

A&gt;Note: F5E19\_190

Query Match 13.24; Score 777.5; DB 2; Length 1123;  
Best Local Similarity 24.04; Pred. No. 2.7e-49;  
Matches 289; Conservative 209; Mismatches 435; Indels 269; Gaps 42;  
QY 38 LVQPGDPKIVRTLVAOCLVCMHNGSQPPADLSFHQVS--SLKELVARVVQRL---CERN 92  
Db 60 LIRDDDIHVRKHLHRCFVVLH--EQTPL-LDPSFWSQREIVERIEMQSGCDC- 115  
QY 93 ERNYLAFGFELLNEARGPPMAFTSSVRSYLPNTVJETLRVSGAMLLLSRVGDDLLVYL 152  
Db 116 -QNVICARYD-----KYDQSSPILELTSSWEFLKRVGHDMVYL 156  
QY 153 LAHCALYLLVPPSCAYVCVCSPLQYICATTDIWPVSASVSRPVRPGRNFTNLRFLQKIK 212  
Db 157 LQQTISIFLLPGKKHQGVSGPL-----CIKHRTLSVHEN----- 192  
QY 213 SSSRQEAAPKPLALPSSGCTKRH-LSLTSTSVPSAKARCYVPVRVEEGPHROVLPFSGKS 271  
Db 193 KERDDNVQP-----PTKQWLSSAVDDCPKDDSATITPIVGEDVDQHKRKTTKRSRI 246  
QY 272 WVPSPASPEVPTRAKDL-----SSKGVS-----DLSLSSVCCKKPKSSTSL 316  
Db 247 YLKRKRQKQKVNFKKVDNAPCITPTSTNGKVSSTGNDENMLHIGINGSL-----TDFV 298  
QY 317 SPQRNAFQLRPFTETRHFLYSRGDCQERLNPFLLSNLQPNLTGARRLVEIIF----- 370  
Db 299 K-----QAKVKENKFKGLSEIVSVIPPHILKTLRPNCSKLLMNHIFGEVNVW 351  
QY 371 -----LGSPPRTSGPLCRTHRLSRRYQWRMLPQOLLVNHAECCQVYLLRSHC----- 418  
Db 352 STTPSHGKGNCP--SGSICLYHSLK--SLKNLIGKTKSHLK-----MLLDKHCVPVLL 402  
QY 419 -----RPTANQOVTALNTSPH-----LMDLLRLHSSPWQVYGLR 456  
Db 403 QEDALKSTTSQSRROKADKL-----PHGSSSQTGPKPCPSVEERKLICTNDQVVSFIW 458  
QY 457 ACLCKVWSASLWGTNRHNERFFKNLKFISLGKYGKLSLOELMMKMKVDE----- 506  
Db 459 AICRYIVPESLGTTHQVLRKNIAFWFSRRNEKCTVQNFLEKVKPSPDPPFPARKELC 518  
QY 507 C-----HMLRSSPGKDRVPAAEHLRLERILAT-FLFWMMDTVYVOLLRSFFYITESTFOKN 561  
Db 519 CMVNGHELQSE-----SIRTOQMLCTKWISLWLEIVKGLVHFNFYATESQGGRL 569  
QY 562 RLFFYRKSWSKLOSIGVROHLERVLRLBELSQEVRHODTWLAMPICRLRFPKPNGLR 621  
Db 570 NIYYRKESWERLISKELSKALDGYVLVDAAEASSRKK-----LSKPRFLPKANGVR 622  
QY 622 PIVNMSYNGTRALGRKQAOHFTQRLKTLFSLMNTYRTKPHLMGSSVLGMNDIYRTWR 681  
Db 623 MYLDFSSSR-----SQSLRDTHAVLKQIQKEPDVGLSSVDFDHDFFYRNLC 669  
QY 682 AFVLRVRLD-QTPRMVYKADVTGAYDAIPQGLKLEVVANMIRHSESTYCIROYAVV-- 738  
Db 670 PYLIHLRSQSGELPPLFVAVADVAFKAFDSVDQGLLHVIOFPLK---DEVILNRCRLVCC 726  
QY 739 -RDS-----QGVHKSFRQVTTLSLQPYMGQFLKLODSASALRNSVVIQSIS 790  
Db 727 GKRSNWNKILVSSDKNSNFSRFTSTV---PY-----NALQ-STVVDKGEN 768  
QY 791 NNESSSSLFDPLHFLRHSHVVKIGDRVCYQCGIPQSSSLTLCSCFCGDMENKL----- 846  
Db 769 HVRKLDLMVWIGNKLNNMLQLODKSFYVQIAGIPQGRHLSLLCCFYVGHLETLIYFP 828  
QY 847 FAEVQD-----GLLLRFVDDFLVTPHLDOAKTFLSTLVHGVPEYGC 889  
Db 829 LEEASKVSKESGREGELIITPSYKLLRFIDYLFVSTSRDQASSFYHLKHFQKDYNC 888  
QY 890 MINLQKTVNFPVEPTLGGAPYQLPAHCLF-----PWCGLLDDTQTLVEFCDYS 940  
Db 889 FMNETKFCINFE-----DKHEHRCSSNRMFVGDNGVPFVRWTLGLLINSRTTEVQVDYT 941

941 GYATSTIKTSLTQSVFVAKTWRNKLVLRLKCHGLFDLQVNSLQVTCVINYKIFLL 1000  
 942 RYLSGHISSTFSAWQKPVNRLQKCYFLVPKCHPILFDSNINSGEIVRLNIYQIFLL 1001  
 1001 QAYRFHACVIQLPDORVRKNLTF-FLGIISQASCCYAILKVPNGMTLKASGSFPP-- 1057  
 1002 AAKFHCYVEVSFRFKLHPQTLFKITI-----SVYMFRLINRVRINTGSSFRVL 1056  
 1058 ----EAAHWLCYQAFLLKLAHSVYIKCILLGPTAQKLLCRKLPEATMTILKAADPAL 1113  
 1057 KLYKEEIVMLGLDAY:QVLKKNSRYRMILLIYKSA--LSKHSLSQQLSSSELRYATDRSN 1114  
 1114 ST 1115  
 1115 SS 1116

SULT 3  
 3838  
 lomerase catalytic chain - fission yeast (Schizosaccharomyces pombe)  
 Alternate names: telomerase reverse transcriptase 1  
 Species: Schizosaccharomyces pombe  
 Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 01-Dec-2000  
 Accession: T03838; T03839; T40085  
 Nakamura, T.M.; Morin, G.B.; Chapman, K.B.; Weinrich, S.L.; Andrews, W.H.; Lingner, J.  
 ience 277, 955-959, 1997  
 Title: Telomerase catalytic subunit homologs from fission yeast and human.  
 Reference number: T15111; MUID:97400623; PMID:9252327  
 Accession: T03838  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-989 <NAK>  
 Cross-references: EMBL:AF015783; NID:92340167; PIDN:AAC49803.1; PID:G2340169  
 Experimental source: strain 972h(-)  
 Accession: T03839  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-524,526-989 <NA2>  
 Cross-references: EMBL:AF015783; NID:92340167; PIDN:AAC49802.1; PID:G2340168  
 Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
 blitted to the EMBL Data Library, March 1998  
 Reference number: Z21904  
 Accession: T40085  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-524,526-989 <LYN>  
 Cross-references: EMBL:AL022299; PIDN:CAAL8391.1; GSPDB:GN00067; SPDB:SPBC29A3.14c  
 Experimental source: strain 972h-; cosmid c29A3  
 Genetics:  
 Map position: 2  
 Introns: 86/3; 113/3; 153/2; 241/1; 372/1; 395/3; 485/3; 524/3; 582/2; 644/1; 693/3; 7  
 Keywords: alternative splicing

Query Match 10.8%; Score 624; DB 2; Length 989;  
 Best Local Similarity 23.7%; Pred. No. 6e-38;  
 Matches 255; Conservative 194; Mismatches 381; Indels 244; Gaps 46;

76 SLKELVARVVQRLCERN---ERNVLAFFGELNE-ARGGPPMAFTSSVRSYLPNTVIELT 131  
 83 SOSELIANVVQKPFDSFERRNLLMKGSMKHEDFRAMVNGVQNDLVSTFPNVLISIL 142  
 132 RVSGAMWLLSRVGDLLVYLAHCAALYLLVPSCAYQVCGSLYQICATTDIWPVSAS 191  
 143 E-SKNWOLLLEIGSDAMHYLLSKGIFEPALPNDNYLIQISGIPLFK----- 187  
 192 YRTPRPVGRNFTNRLFLQIKSSROEAPKPLPSRGTKRHLSTTSV-----PS 243  
 188 -----NNV--FEETVSKKRTTETITQKSARKEVSNWSISISRFSIFRYS 234  
 244 AKKARCVVPRVEGPHRQVLPTPFGKSWVPSPARSPEVPTAEKOLSSKGKVSILS-G 302

235 YKFK-----QDLYFNLIH 247  
 303 SYCCKHPESTSLSPPRQ---NAFQLRPFIETRHFLYSGDQERLNPFLSLNLPN 358  
 248 SICDRNTVHMWLQWIFPRQFGLINAFQVKL-----HKVPLVQSQTVP- 292  
 359 LTGARLVEIIFLGRSPRTSGPLCETHRLSRRYQWMBPLFQOLLVNHAEQYVLLRSHC 418  
 293 ----KLLKY-----PLIETAKRLHRIS-----LSKYNH-YCFYI- 325  
 419 RERTANQOVTDALNTSPPELMDLLRLHSSPWQVYGFRLACICKVVASLWGRNRRRF- 477  
 326 ----DTHDDE-----KILSYSLKPNQVAFRLSILVRVFPKLING---NQRIFE 367  
 478 --FKNLKXFIISLKYKLSLOELMWKMKVEDCHWL-----RSSPGKDRVPAAEHLRERIL 531  
 368 IILKOLETFLKLSRYESFSLHLYMSNIKISIEIWLVLGKRSNA---KMLSDFEKRIQIF 424  
 532 ATFLFWMIDTVVQLLRSPFYITSTFQKNLRFYRKSVMSKLSQISGVROHLERVRL--- 588  
 425 AEFYWLNSFIIPILQSFYITSSDLRNTVFRKDIWKL-----CRFFITSMKEAF 480  
 589 RELSQEVRHHQDTW-LAMPICRLRFIPKPNGLRPIVMSYSMTGTRALGRRKQAOHT-Q 646  
 481 EKINENNRV--MDTQKTLPPAVIRLLPKKNTFRLLITNLKRFLIKQMGSKMLVSTNQ 538  
 647 RLKTLFSLNRYERTKPHLMG--SSVLGMN-DIYR---TWRAFVLVRALDQTPMYFVK 700  
 539 TLRPVASILK-----HLNEESSGIPFNLEVYMKULTFKKOLLKRMFGR--KTYFVR 589  
 701 ADVTGAYDAIPQGLVEVVMNIRHSESTYCIROYAVVRDSDQGVHKSFRQVTTLSDL 760  
 590 IDIKSCYDRIKQDLMFRIVKVKLXDPE--FVIRKYATIHATSD--RATKNFVSEAFSYFDM 646  
 761 QPY--MGQELKHLQSDASALNSVIVQSISSMNESSSLFDFFLHFLRHSVYKIGRCY 818  
 647 VPFEKVQVLLS-MKTS-----TLFVDFVDYWTKSSSEIFKMLKHLSGHIVKIGNSQY 699  
 819 TQCGIPOGSSLSLTLSCFCGDMENKLFABVQDQDGLLLRFVDDFLVTPHLDQAATFL 877  
 700 LQKVGIPQGSILSSFLCHFYMEDLIDEVLSFTKKKGVLLRVVDDFLFIIVNKKDAKKFL 759  
 878 STLHGVVPYGCINLQKTVNVPVEPGLTGAAPYQLPAHCLFPWCGLLDQTLTLEVF- 936  
 760 NLSLRGFQKHNFSTSLKTVINFENSNGIINNT--FFNESKKRMPPFGFSVNMESLDTLL 817  
 937 ---CDYSGYAQTSIK-TSLTFQSVFKAGTMRNKLJVLRLKCHGLFDLQVNSLQVTC 991  
 818 ACPKIDEALFNSTVELTKHMKGSFFY--KILRSSLASFAQ-----VFIDITHNSKFNSC 870  
 992 INIYKI-----FLQAYRFHACVIQLP-----FPQVRKNLTFFLGISSQASC 1035  
 871 CNIYELGYSMCVRQAAYLKRMDKFIQPMFEITDLNVIQKIKWKLAEILGYTSRRFLS 930  
 1036 CYAILKVKNPQNT--LKASGSFPPPEAAHMLCYQAFLLKLAHSHVYIKCLLGPLR 1087  
 931 SAEVKNLFCNLGRDLKPFKYPH-----CFEQLTYQFQSLTDLIK-----PLR 974

RESULT 4  
 T31107  
 telomerase reverse transcriptase - Oxytricha trifallax  
 C;Species: Oxytricha trifallax  
 C;date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
 C;Accession: T31107  
 R;Bryan, T.M.; Sperger, J.M.; Chapman, K.B.; Cech, T.R.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 8479-8484, 1998  
 A;Title: Telomerase reverse transcriptase genes in Tetrahymena thermophila and Oxytricha  
 A;Reference number: Z20985; MUID:98337940; PMID:9671703  
 A;Accession: T31107  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-1132 <BRY>

;Cross-references: EMBL:AF060230; NID:g3342795; PID:g3342796; PIDN:AAC39163.1  
;Genetics:  
;Gene: TERT

```
Query Match      7.2%; Score 426.5; DB 2; Length 1132;  
Best Local Similarity 19.4%; Pred. No. 3.8e-23;  
Matches 194; Conservative 204; Mismatches 418; Indels 185; Gaps 37;  
  
y 75 SSKELVARVORLCERNVNL-AFGELLNARGGP---PMAFTSSVRSYLPNTVI-- 128  
b 153 NYIWMISKVIELLIENKLNKLTGFKLVNNEFGNQHLMMQOQDSSHNSNEFWKC 212  
  
y 129 -----ETLRVGAWMLLSRVGDDLLVYLAHACALYLLVPPSCAYOVCGSPVQICATT 182  
b 213 DYINLNKQCMITKWEKYFYLDHFLPHIYKEYWIFLKRDESLVQISGNTIP--CYLN 270  
  
y 183 DTWPSVASRYRTRVGRNFTNRLFLQOIKSSQBEAPKPLALPSRTKXHLSTSTSV 242  
b 271 EKLGRLOAAAF--YEGFNKAAN-----SAAQCSNPEANDLISAPQRK---INTAI- 315  
  
y 243 SAKKARCYVPVRVSGPHRQVLPTPSGKSWVPSPARSPVPTAEKDLSSGKYSD-LSL 301  
b 316 VMKTKHYNKKADE-----SYLTN-----QEGFWDQDKRN 348  
  
y 302 GSVCKKHPSSTLSLSPRONAFQLRPFIETRHLYSRGDQERLNPFLSLNLPNLTG 361  
b 349 RLIFYCAH-----QNR-----FQGHLSNKTLSQQQIR-----DNRYKEVFG 385  
  
y 362 ASRLVEIIFLGSRPRTSGPLCETHLSRRYQWREPLFOOLLVNHAEQYVRLRSHC--- 418  
b 386 FRNV-----RSLKG-----KWSIIEQVIVNOKKDFKYLSKNCPLP 424  
  
y 419 -RFTANQOVTDALNTS---PPHMLDLRLHSSPWQVYGLRACLCKVVSASLWGRHN 473  
b 425 ENWKNLKSFLEDAAVSGELRQVFRQLFECQDQORQISNLTFFVANVPKFNLEK-N 483  
  
y 474 ERFPKULKFTSLGKYKLSLQELMWKQVEDCHWLRSSEPGKORVPAAEHRLERILAT 533  
b 484 KKFINKMQLQFKVRFESFTKISILNFRYNEVSWL-SFKCDENKFFMNEHVFVK 542  
  
y 534 FLFLMDTVVOLLRSFFVITESTQKRLFFYKSVMSKLSQISGVQHLERVLRLS 593  
b 543 VLKWPEDLAITLMKCYFISTEKAKEYQRIEYRNKNIMRLSIDDLKQ-NLKQVEK 601  
  
y 594 EVRRHQDTWLAMPICRLRIFPKPNGLRPIVNMYSVMGTRALGRKQAOHRTORLTKFS 653  
b 602 KEMRIFCESQNPAP-GKRLIPKGTFRPIMTFNRKIPNQ-VGKFSQRMITNNKLQTAHM 659  
  
y 654 ML-NYERTKPHLMGSSVLGMNDIVRTWRAFLVRALDQTPRMVYFKADVTGAYDAIPQ 712  
b 660 MLKXLSKMKFKSFGFAVFNVDIMKRYENFVQWKQIN-SPKUIFYVAMDIKCYDNVD- 717  
  
y 713 GKLEVEVANMRHS---ESTYCIROYAVVRDSOGVHKS-FRQVTTLSDLQPYMGQFL 768  
b 718 ---CERVNFKQSDLMQKEYFILNTEVLKKNNIIVERSNFRK---LPKQYFYKFKQ 770  
  
y 769 K-----HLQD---SDASALNSVIEQSI SMTESSSSSLPDPFLFLHSHSVKI 813  
b 771 KIGIDSSYPTTFEILEDEFDNLN-MKRTIIVBQEQRKFPKNDLLOVFLKICNNXVTF 829  
  
y 814 GDRCTQCQIPQSGSLSTLLCSLCFGDMENKLPFAEVORD-----GILLRFVDDFL 864  
b 830 NKKQYQMKGIPOGLCVSVILSSFYANLEENALQFLAKESMDPEKEINLLMRLLTDYL 889  
  
y 865 LVTHPLDQAKTFLSTLVHGVBEYGMINLQKTVNFPVEPOTLGGAAPOYLPDHC----- 919  
b 890 LMTTEKNAMVLEIEKLYQLSGNFFKFMKMLKTNFALNLQIG-----CTNTTQ 939  
  
y 920 -----LFPWCGLLDTQLEVFCDYS---GVAOTSIKTSITFOSVFKAGTWENKL 967  
b 940 DIDSINDDLFWHIGISIDIKTLNIQININIKSEIGLICTLVNVMGTNESIILWKLKSL 999  
  
y 968 LSVLRKCHGLFLDLQVNSLQTVNCINIKYIFLQAYRFHAC 1008
```

Db 1000 MNNT-----SFYFKSTINTKQFANITLSKLYIAABEKYVAC 1035

## RESULT 5

T14891  
telomerase (EC 2.7.7.-) catalytic chain p133 - Tetrahymena thermophila  
N:Alternate names: telomerase reverse transcriptase  
C:Species: Tetrahymena thermophila  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 11-May-2000  
C:Accession: T14891  
R:Collins, K.; Gandhi, L.  
Proc. Natl. Acad. Sci. U.S.A. 95, 8485-8490, 1998  
A:Title: The reverse transcriptase component of the Tetrahymena telomerase ribonucleoprotein complex  
A:Reference number: 218252; MUID:98337941; PMID:9671704  
A:Accession: T14891  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1117 <COL>  
A:Cross-references: EMBL:AF061284; NID:g3335166; PID:g3335167; PIDN:AAC39140.1  
C:Genetics:  
A:Gene: TERT  
A:Genetic code: SCS  
C:Keywords: nucleotidyltransferase

Query Match 6.8%; Score 401.5; DB 2; Length 1117;  
Best Local Similarity 21.2%; Pred. No. 2.7e-21;  
Matches 173; Conservative 147; Mismatches 358; Indels 139; Gaps 28;

```
Qy 296 SDSLGSVCCKKPSSTLSLSPRONAFQLRPFIETRHLYSRGDQF----- 344  
Db 190 NFNMGKATSNNNNNANLSNEKQENQYIPEIQRSQIFCYNHMGREPGVFKSFFNY 249  
  
Qy 345 -RLNPSFLLSNLPNTGARRLVEIIFLGSRPRT--SGPLCRTHLSRRY--WQMRPLF- 398  
Db 250 SEIKKGQFQVIEKLGROQFINSDKIKPDHPOTIKTKLLKEYQ-SKNFSQBEEDLEL 308  
  
Qy 399 ---OQLVNHAEQYVRLRSHCRF-----RTANQOVTDALNTSPH 437  
Db 309 EFTKIVQNFHNINFNLYLLKFKCLPENTQSLKSQVKQIVQSENKANQOCSNLFNS--- 365  
  
Qy 438 LMDLLRLHSSPWQVYGLRACLCKVVSASLWGRHNERFFKNLKKFISLGVKLSLQ 497  
Db 366 ---LYDTEISYKQITNFIQIIONCVPNQLG-KKNFKVLEKLYEFQMKFENQKVD 421  
  
Qy 498 LWMKMYVEDCHWLRSSEPGKORVPAABH-RLRERILATFLFLMDTVVOLLRSFFYTES 556  
Db 422 YICFMDVDFVEWFVDLKNQKFTQKRYISDKRILGDLIVFIINKIVIPVLYNFYITEK 481  
  
Qy 557 TFQKRLFFYKRSVMSKLSQISGVQHLERVLRLSQQEVRHH---QDTWLAMPICRLRP 613  
Db 482 HKGSQ:FYKRPFWLKVSKLTI-----VKLEENLEKVEKLPEDSFQYPOGKURI 535  
  
Qy 614 IPKPNGLRPIVNMYSVMGTRALGRKQAOHFTQRLKTLF--SMLAYERTKH--PHLMGSS 669  
Db 536 IPKGSFRPIWTL-----RKDKQKNILNQLIILMDLSQLVFNLMGLGQKIGVS 586  
  
Qy 670 VLGMNDIYRTWRAFLVRALDQTPRMVYFKADVTGAYDAIPQGLKLEVEVANMRHS--- 726  
Db 587 VFDNKQISEFAQFIEKWNKGR-PQLYVYVTLDIKKCYDSIDQMKLL-----NFFNQSDLI 641  
  
Qy 727 ESTYCIROYAVVRDSOGVHKSFRQVTTLSLQ-----PYM-----G 765  
Db 642 QDTYFINKYLLFOENKRPLLQIOQTNLNSAMEIBEKINKKPFKMDNINFFYFNLER 701  
  
Qy 766 QFLKHLQSDASAL-----RNSVVISQSI SMTESSSSSLPDPFLFLHSHSVKIG 814  
Db 702 QIAYSLYDDDDQILQKGFKEIQSDRPFVINOQKPRCITKDIHNLKHIQYVNIQSFN 761  
  
Qy 815 DRCYTCQGIPOGSSSLTLLCSLCFGDMENK---LFAEVQDQ---LLLRVDDFLVLT 867  
Db 762 KVPKQKRGIPQGLNLSVGLCSFYFKLHEEYQFLKNAEQVNGSINLLMRLLTDYLFIS 821
```

868 PHLDQAKTFLSTLVHGVPEYGCMLNOKTVN--FPVEPOTLGAAPYQPAHCLPFWCG 925  
822 DSQQNALNLIQLOCANNGFMFDQKITTFQFQEDYNL---BHFKISVQNECQWIG 878  
926 LLLDPTQTEVCDYSGYAQTSIKTSLTFQSVFKAGKMRNKLKSLVLRKCHGLFLDL--- 982  
879 KSIDNWTLEI-----KSIQKQTOQEQINQINVAISIKN-LKSQKXNKLRSFLNQLID 930  
983 ----QVNSLQVNCINYY---KIFLLQAYRFHACVIQL 1012  
931 YFNPNNINSFEGLCRLQYHRSKATVMKFPFMTKLFOI 967  
RESULT 6  
53396  
;Alomerase catalytic chain EST2 - yeast (Saccharomyces cerevisiae)  
;Alternate names: protein U843.12; protein YLR318w  
;Species: Saccharomyces cerevisiae  
;Date: 05-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 05-Nov-1999  
;Accession: S53396  
;Du, Z.  
;Submitted to the EMBL Data Library, February 1995  
;Description: The sequence of S. cerevisiae cosmid 8543.  
;Reference number: S53390  
;Accession: S53396  
;Molecule type: DNA  
;Residues: 1-884 <DU2>  
;Cross-references: EMBL:U20618; NID:92258165; PID:96621136; GSPDB:GN00012; MIPS:YLR318w  
;Genetics:  
;Experimental source: strain S288C (AB972)  
;Gene: SGD:EST2; MIPS:YLR318w  
;Cross-references: SGD:S0004310; MIPS:YLR318w  
;Map position: 12R  
Query Match 5.8%; Score 341; DB 2; Length 884;  
Best Local Similarity 19.9%; Pred. No. 6.1e-17;  
Matches 206; Conservative 147; Mismatches 355; Indels 326; Gaps 39;  
48 RVLQAQCLVCMHWSQPPADLSFHQVSSSLKELVARVVQRLCERNERNVLAFGELL-NE 106  
48 RXIALPCL-----PGDLSHKAV-----IDHCIIYLLTGELYNVLTFGYKARNE 92  
107 ARGGPMAFTSVRSYLPNTVETLURVSGAWMLLSRVGDDLLVYLAHALYLLPPSPC 166  
93 DYNSLGFCHSANV-----VTLLKGAAMKMFSLVGTVAFDLLIN---YTVI--- 137  
167 AYQVCGSPLOYI---CATTDIWP-----SVGASYRTPVGRNFTNLRFFQQIKSSS 215  
138 ---QFNGQFTQIVGNRCNEPPLPKXWVORSSSSATAAQIKQLTEPTVTKQFLHK--- 190  
216 ROEAPKPLALPSRGTKRHLSTLSTSVPSAKKARCYPVPRVEEGPHRQVLPSPGKSWVPS 275  
191 -----LNINSSSF-----PPYSKILPSS----- 208  
276 PARSPEVPTAEKLSKGVSDLSLGSVCCKHKPSSSTLSLSPRONAFQLRPFIETRHF 335  
209 -----SSIKKLTDL-----REAIFFPNLVKIP----- 230  
336 LYSRGDDGOERLNPSSLNLOPNLTGARRLVEIIFLGSRRPTSGPLCRTHLSRRYQWR 395  
231 -----QRL-----KVRINLT----- 240  
396 PLFQQLLVNHABCOVYLLRSHCRPTANQOVTALNTSP---HLMDLRL---HSPWQ 450  
241 ---LQKLLKRLKELNVTSLNSIC-----PPELGTVLDLSHLRSQSPKER 282  
451 VYGFRACLCKVVSASLWGRTHNEREFKNLKFKISLGKVGKLSQLMKMKVVEDCHWL 510  
283 VLKFIIVILQKLLPOEMFGSKNKGKIKNLNLLSLPLNGYLPFDSLLKJLKRFRWL 342  
511 RSSPGKORVPAAEHLR---ERLATFLFWMMDTVYVQLLRSFFVITE---STFOKNRLFFY 566  
343 FTSD-----IWFTHKNFENLNQALICFISWLPQLPKTIQTFYCYTEISSTV---TIVVF 395

567 RKSWSKLSIGVROHLERVRLRELSQBEV---RHQDQTLWAMPICLRPIPKP---NGLR- 621  
396 RHDWNKLIPTFFIVEFKTY---LVENNVCNRHNSYTLNFNHNSKMRIPKKSNEPRI 451  
622 ---PIVNWYSMTGTRALGRKQQAQHFQRLKTLFSLNLYERTKPHLMGSSVLGKNDIYR 678  
452 IAPCRGADEEETIYENHKNAIQPTQKI-----LEYLRNKRPTSP-TKIYSTQTAD 504  
679 TWRAFVLV---RALDQTPRMVYFKADVTGAYDAIPQGLVFEVVANNIRHSESTYCIROY- 735  
505 RIKEFKQRLKKFNVLPELYFMKFDVKSCYDISIPRMECMRILKDALKNENGFFVRSQYF 564  
736 -----AVVRDSQCVHKSFRROVTTLSLDLPYMGQFLKHLQSDASALRNVV 784  
565 FNTWTVGLKLFNVNVSARVPKPYELIDNVRTV-----HLSNQDV----- 604  
785 IEQISMNNESSSLFDFFLHFLRHSVKIGDCRYTCQOIGIQGSSLSLTLCLSCFQDM-- 842  
605 -----INVMEIIF-----KTALWVEDKCIYREDGLFGQSSLSAPIVDLVYDLE 650  
843 -ENKLFRAEVDGILLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGCMLNOKTVNPP 901  
651 FYSEFKASPSQDTLILKLADDFLIISTDOQOVINIKKLAWGQFQKYNANRDKIL---- 706  
902 VEPCTLGGAAPYQLPAAHCLPFWCGLLDQTLEVFCDYSGYAQTSIKTSLTFQSVFKAGK 961  
707 -----AVSSQSDDDTVIQFCAMHIFVKELEVWKHS-----TWNNFHRSKSK 750  
962 TMRNKLVLRLKCHGLFLDLQVNSLQTVGINI-----YK-----IFLLQAYR 1004  
751 GIPRSLIALFNTRISYKTIIDTNLNTNTVLMQIDHVVKNISECYKSAFADLSINVTQNMQ 810  
1005 FHACVQLPFDQV 1018  
811 FHS-----FLQRI 818  
RESULT 7  
T20392  
hypothetical protein DY3.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T20392  
R:Lennard, N.  
submitted to the EMBL Data Library, June 1997  
A:Reference number: Z19266  
A:Accession: T20392  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-561 <WIL>  
A:Cross-references: EMBL:Z96047; PIDN:CAB09413.1; GSPDB:GN00019; CESP:DY3.4  
A:Experimental source: clone DY3  
C:Genetics:  
A:Gene: CESP:DY3.4  
A:Map position: 1  
A:Introns: 44/3; 106/2; 164/1; 249/3; 306/1; 357/2; 409/3; 495/3; 517/3  
Query Match 2.3%; Score 133.5; DB 2; Length 561;  
Best Local Similarity 19.7%; Pred. No. 0.089;  
Matches 66; Conservative 65; Mismatches 149; Indels 55; Gaps 12;  
617 PNGLRPIV---NMSYS-MGTRALGRKQQAQHFQRL---KTLFSLNLY---ERT-KHPHLMG 667  
249 PHFRPNVATFKLISISQKLRPLFKKKAIDKKTETQWKRLNSMLSWCLERSGVVHTIR 308  
668 SSVLGMDIYRTWRAFVLRVRLDQTPRMVYFKADVTGAYDAIPQGLVFEVVANNIRHSE 727  
309 DSKKVSDF-----LKKNSQNSKIIGYTADYSKCFSTVNDVLSIIDRPSQSH 358  
728 STYCIROYAVVRDSQCVHKSFRROVTTLSLDLPYMGQFLKHLQSDASALRNVVIEQ 787  
359 DIYTV-----CGKGRNHGGPHKLFCSAGTELNAHEALR---RKVELKGVNEFVCYR 408

```
Y 788 STSMNESSSLPFDLHFLHRSVWVGKIDRCYTCQCGIPQSSSLTLCSCFCGDMENKLP 847
b 409 EMS-----SSTLLSVIRTLSTLYYKRGFTSWRTKGVPOGHPISSNLAHMYLNNFEQKYW 465
Y 848 ABVQDGLLL--RFVDDFLVTPHLDQAKTFLSTLVHGVPYGCMLNLOKTVVNFVPEPG 905
b 466 SNEKEDSRVFCRYEDDDIFITFENSLEKMKELSTGNTHFTANPKK-----515
Y 906 TLGGAAPVLPAC-----LFPWCGLLLTQTLFV 936
b 516 -----FKKSRCCGASQVLQWCGVKLDFQSGNCF 543

RESULT 8
98119
;Species: Streptococcus pneumoniae (strain R
;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
;Accession: E98119
;Host: Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; E
; R.; Leslanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
; P.; Sun, P.M.; Winkler, M.E.
; Bacteriol. 183, 5709-5717, 2001
; Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
; Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
; Reference number: A97872; MUID:21429245; PMID:11544234
; Accession: E98119
; Status: preliminary
; Molecule type: DNA
; Residues: 1-330 <KUR>
; Cross-references: GB:AE007317; PIDN:AAL00786.1; PID:G15459687; GSPDB:GN00174
; Genetics:
; Gene: transposase H

Query Match 2.2%; Score 131.5; DB 2; Length 330;
Best Local Similarity 19.4%; Pred. No. 0.057;
Matches 59; Conservative 34; Mismatches 80; Indels 131; Gaps 8;

607 PICRLFIKPKGLRPIVNMYSMGTALGRKKAQHFTQRLKTLFSLMNYERTKPHLM 666
b 3 PICEPHSDTSYGFRL--NRSC-----KAIMKLEY-----32

667 GSSVLGNNDIYRTWAFVLRVLRALDQTPRMVYKADVTGAYDAIPQGLKVEVVMNTRHS 726
b 33 -----LNDGYE-----MTVIDLEKFFDTVPQDRMLSLVHNIIDEG 68

727 ESTYCIROYAVVRDQGVHKSFRQVTTLSLQPYMGQFLKHLQDSASALRNSVVE 786
b 69 DTESLRKYL-----78

787 QSISMNESSSLPFDLHFLHRSVWVGKIDRCYTCQCGIPQSSSLTLCSCFCGDMENKL 846
b 79 -----HSGVIINGCRYKTLVGTVPQGNLSPLLSNI-----MLNEL 113

847 FAEVQRDGL--LRFVDDFLVTPHLDQAKTFLSTLVHGVPYGCMLNLOKTVVNFVPEP 904
b 114 DKEELKRLGRFVRYADDCVITVVSEAAKRMYSVSRFIEKRLGLKLVNNTKTLTRPREL 173
b 905 GTLG 908
b 174 KYLG 177

RESULT 9
38487
;Species: Homo sapiens (man)
;Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 05-Nov-1999
;Accession: I38487
; Fukuda, M.N.; Sato, T.; Nakayama, J.; Klier, G.; Mikami, M.; Aoki, D.; Nozawa, S.
; Genes Dev. 9, 1199-1210, 1995
; Title: Trophinin and tastin, a novel cell adhesion molecule complex with potential inv
```

```
A;Reference number: I38487; MUID:95278733; PMID:7758945
A;Accession: I38487
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-778 <RES>
A;Cross-references: EMBL:U04810; NID:g905355; PIDN:AAA79333.1; PID:g905356

Query Match 2.2%; Score 131; DB 2; Length 778;
Best Local Similarity 22.4%; Pred. No. 0.22;
Matches 107; Conservative 44; Mismatches 167; Indels 160; Gaps 26;

QY 2 TRAPRCFAVRSLLRSRYREVMPLATFVRRLGPEGRRLVQGPDKYIRTLVA-----QC 54
Db 358 TPEPSTPRVQ-----AQLRGVSPQS-----CSEDPALPWEQVAVRLFQDESC 401

QY 55 LVCMHWSQPP---PADLSFHOYSSILKELVAR---VVORLCERNERNVLAFGFELNNEAR 108
Db 402 IRSLEGSGKPPVATPSPGPHNSRTPSQEVKIQIGILQQLLQKQEVGLVGGQCVPLN---458

QY 109 GGP--PMAFTSSVRSYLPNTVIETLRVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSC 166
Db 459 GGSILDMVELQPLLTETISRTLNATEHNSGT-----SHLPGLLKHSL-----PKFC 504

QY 167 AYOVCGSLVQICATTDIMP-----SYSASRPTRPVGRNFTN-L 205
Db 505 LPFECGEP--QPCFPAPGPPFAFCRSEPEIPEPSLQEQLEVPEPYPPAPB--RPLESCC 560

QY 206 RFLQIKSSRQE-----APKPL-----ALPSRGTKRHLSTLTSVPSAKK 246
Db 561 RSEPEIPESRQEQLEVPEPCPAERPLESYCRIFEIPESSRQQL-----EYPE---612

QY 247 ARCVPVPRVEGPHRQVLPTPSGKSWVPSPARSPVPTAKDLSSGKVSDDLSSGSV--304
Db 613 ----PCPPAEPGP---LQPSTQCGSGPPGPC-----PRVELGASEPCTLEHRSLESSLP 660

QY 305 CCKH-KQSSTLL-----SPPRQNAFLAPFFIETRHFLYSRGDDGGERLNPSLLGN 354
Db 661 CCSQWAGATTLIFSSQHPICASPPICSLSQSLAP-----PAGQAG-----LSN 703

QY 355 LQPNLTGARRIV-----EIIFLGSRPRTSGP--LCRTHRLSRRYVQ 393
Db 704 LAPRTLALRESLKLCLTAIHCHEARLDDCAFTVTSRASPSPGTRVCTNPVATLLEWQ 761

RESULT 10
T24227
hypothetical protein R166.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T24227
R;Matthews, P.
submitted to the EMBL Data Library, August 1995
A;Reference number: Z19859
A;Accession: T24227
A;Status: preliminary; translated from GB/EMBL/DD8J
A;Molecule type: DNA
A;Residues: 1-502 <WIL>
A;Cross-references: EMBL:Z50795; PIDN:CAA90562.1; GSPDB:GN00020; CESP:R166.1
C;Genetics:
A;Gene: CESP:R166.1
A;Map position: 2
A;Introns: 32/1; 164/1; 306/1; 351/2; 480/3

Query Match 2.2%; Score 129; DB 2; Length 502;
Best Local Similarity 22.0%; Pred. No. 0.16;
Matches 100; Conservative 57; Mismatches 159; Indels 138; Gaps 22;

QY 141 LSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPLYQICATTD----IMPSVSASVRRPFR 196
Db 70 LAEYSDQDTAFNLA--ALQIQGPP-----PPLNYPAGTDFMALLPGLIAATSP--117
QY 197 PVGRNFTNRLFLOQIKSSROEAPKPLALPSRGTKRHLSTLTSVPSAKKARCYVPRVE 256
```

118 -----KFPRLRLSLSSVAEKVSTPSEAAATSSSSPSLNLTSSNSVPLAKFPSTPLE 173  
257 E-----GPHRQVLPTPSGSKSWVPS-----PARSEVPTAEKDLSSK-----G 293  
174 SLASDQSSSTSSVRSVLPSTSSHTSHPAGILPATTTNVAAPVPPSSRATANVFG 233  
294 KVSDDLSSGSV-CCKH--KPSSTSLSP-----RQNAFLRPPTETRHFLYS 338  
234 NSIGLNFSGAASVTRHLVVPSSSTISQOFTSGRSSSTGQKEGSSSPFLGVG--YS 290  
339 RGDQERLNPFLSLNQPNLTGARRLVEIFLGSRPRTSGPLCRTHLSRRYQWRPLF 398  
291 QPYGNDVFS--LGDFTDN-----NPSLTENPTLSAQISRL----- 324  
399 QOLLVNHAEQVY-----RLLRSHCRFTANCOVTDALNTSP--PHLMDLLRLHS 446  
325 -----AECALAAKSNLPLPLPVNKRR--VSKEVIELLKCFATPSMIHAFKYS 374  
447 SPWQVYGF-----RACLKVVVSASLWGTNRNRRFFKNLKKFIS 486  
375 A--IYGRFDTKRPHKVLTILHETTVAEAQCLLVPSLL--TRDE--LPLARQIVK 427  
487 LGKYGKLSQLQELMMKMKVEDCHLRSFGKDRVP 520  
428 DAGNYAKSKR--PCDPAHLHSPISPGNSPPP 459

RESULT 11  
22710  
VA-directed DNA polymerase homolog T914.6 - Arabidopsis thaliana  
Alternate names: hypothetical protein At2g28980  
Species: Arabidopsis thaliana (mouse-ear cress)  
Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Feb-2001  
Accession: T02730; C84691  
Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Roun-  
abmitted to the EMBL Data Library, August 1998  
Description: Arabidopsis thaliana chromosome II BAC T914 genomic sequence.  
Reference number: Z14710  
Accession: T02730  
Status: translated from GB/EMBL/DBJ  
Molecule type: DNA  
Residues: 1-1529 <ROU>  
Cross-references: EMBL:AC005315; NID:g3461834; PID:g3461840  
Experimental source: cultivar Columbia  
Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.  
iss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
ature 402, 761-768, 1999  
Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
Reference number: A84420; MUID:20083487; PMID:10617197  
Accession: C84691  
Status: Preliminary  
Molecule type: DNA  
Residues: 1-1529 <STO>  
Cross-references: GB:AE002093; NID:g3461840; PIDN:AAC33226.1; GSPDB:GN00139  
Genetics:  
Gene: At2g28980; T914.6  
Map position: 2  
Introns: 428/1; 532/3

Query Match 2.1k; Score 125; DB 2; Length 1529;  
Best Local Similarity 19.2k; Pred No. 1.7; 341; Indels 394; Gaps 58;  
Matches 209; Conservative 147; Mismatches 147; Indels 394; Gaps 58;  
142 SRVGDDLLVYLAHALCALYLAVPPSCAYQVCGSPYQICATTDPWPSVASYAPTRPVGRN 201  
220 SETGVDAWVEY-----KYPWLPFRCSGCKWGHQEVCLTR--PSPNQLSPT----- 265  
202 FTNLRFLLQIKSSSRQEP---KPLALPSGTRKHLSTLTSVPSAKKACYPV--PRV 255  
266 -----ELETEDKTEPPLMKKPEILUSKSPSATLTKTLNGDSHTQKV---PMKNPTV 314

RESULT 12  
T30838  
cytoplasmic dynein heavy chain - Paramesium tetraurelia  
C;Species: Paramesium tetraurelia

QY 256 -----BEGPHRQVLPTPSGSKSWVPSPAR-----SP-----EVPTA 285  
Db 315 LONKGEVAEENENLKDGPWHITVSPAKVGRSQPQOEVVNHNASPRFVLAVEEDTA 374  
QY 286 EKDLSSGKGVSDLSLSSGVCCKHKPSSSTSLSPPRQNAFQLRPFT-----ETRHFLYSRGD 341  
Db 375 DTNGHEEGEIVLSQ-----QLNDNSDSV-----GVQGTPEYIPRISKSHQKVKSSA 422  
QY 342 QGB-----RLNPSFLLNL-----OPNLTGARRLVEI 368  
Db 423 NQNTKGRIVWVRDNARLTVPFKSSOMITCSILLEGKEBEFFCSFIYASNFVEER---I 479  
QY 369 IFLGSRPRTSGPLCRTHLSRRYQWRPLFQOLLVNHAEQVYVRLLRSHCRFTANCOV 428  
Db 480 LWEDIRSHDSPILR-----NKPWILCGDFNEILEGGHSHY----- 516  
QY 429 DALNTGPPHMDLLRLHSPWQVYGFRLACLVKCVVSASLWGTNRNRRFFKNLKKFISLG 488  
Db 517 DNGSPYTPPGWRD-----FOEIGRL-----MLEAAATGR--KPF--KFVNV- 553  
QY 489 KYGKLSQLQELMMKMKVEDCHLRSF-----GKORVPAE 523  
Db 554 -----LTKLPQFLPVVESHNAASSAPLYVSTYALYRFSKKLTKLPHLRELGEKIGDLP 607  
QY 524 HRLRE-RIL-----ATFELWMDTYVQLLRSFFYIT-----ESTFQK----- 560  
Db 608 KRTREAHILLCEKQATTLANPSQETIAELKAYTDWTHLSELEGFLKQSKLHMMVGD 667  
QY 561 -NLPFYKSVWSKLSQIGVROHLERVLREL-----SQBEVRHODTWLAMPICR 610  
Db 668 GNNSYFHKAQVRKM-----RNSIREIRGNPAETLOTSEIKGAEARFFN----- 712  
QY 611 LRPIKPNG-----LRPIVNMVSYNGTRALGRKQAOHQFTQRLKTLFSLMNIYRTK 661  
Db 713 -EFLNCSQDFHGISVEDLRLNLSYRCSVTQDNILTRVTGEEI-Q--KYLAMPN---NK 766  
QY 662 HPHLMGSSVGNNDIYR-TWRA-----FVLVRALDQTPMYTFVADVTGADAIPOGKLIV 716  
Db 767 SPGPDGYT-----SEFFKATWSLTGPDF-AAIQS-----FFVKG-----FLPKGLNA 808  
QY 717 EVVANMIRHSES-----TYCIRQYAVVRDSQGVHKSFRROVTTLSDLQPYMGQF 767  
Db 809 TILALIPKDEAIEMKDVRIPOCCNVLYKVISKILAN-----RLKLLPSFILQNSAF 862  
QY 768 LKHLDQSDASALNSV-----IEQISMNSESSSLFD-----FFLHFL----- 806  
Db 863 VKERLLMENVLLATELVXDYHKESVTPRCAMKIDISKAFDSVQVQFLNTLEALNFPETF 922  
QY 807 RH-----SVVKIGDRC--YTQCGIPQSSSLSTLLCSLCFGDMENKL-FAEVQR 852  
Db 923 RHWIKLCISTATFSVQVNGELAGFFGSSRGLRQGCALSPYLFVICMNVLSHMDIAAVHR 982  
QY 853 D-----GLL-LRFVDDPLL-VTPHLDOAKTFLSTLVHGVPEYGCMLNLOK----- 895  
Db 983 NIGYHPKCEKIGLTHLCFADDLVFPVDGHWISGVINVPKFAGRSGLQISLEKSTIYL 1042  
QY 896 -----TVNPFVBPPTLGGAAFPQHPAHCCLFPWCGLLLDTQTLVEVPCDYGYAQT 945  
Db 1043 AGVSASDRVQTLSSFPFANG-----QLPVRYL-----GLPLITKQMTT-ADYSPLIE- 1088  
QY 946 SIKT-----SUTFQSVFKAGTKWENKLLSVLRKCHGLFLDLQVNSLQTVICINIKFLLQA 1002  
Db 1089 AVKTKLSSWTARSLSYAGR-----LALLNSVVISIAN-FWMSA 1125  
QY 1003 YRFHA-CVIQL 1012  
Db 1126 YRLPAGCIREI 1136

```
;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 02-Mar-2001
;Accession: T30838; PC4341
;Kandl, K.A.; Forney, J.D.; Asai, D.J.
;Submitted to the EMBL Data Library, January 1995
;Description: The dynein genes of Paramecium: the differential expression of axonemal a
;Reference number: Z20502
;Accession: T30838
;Status: Preliminary; translated from GB/EMBL/DBJ
;Molecule type: DNA
;Residues: 1-4540 <KAN>
;Cross-references: EMBL:U020449; NID:G987228; PID:G987229; PIDN:AAA75445.1
;Asai, D.J.; Beckwith, S.M.; Kandl, K.A.; Keating, H.H.; Tjandra, H.; Forney, J.D.
;Cell Sci. 107, 839-847, 1994
;Title: The dynein genes of Paramecium tetraurelia: Sequences adjacent to the catalytic
;Reference number: PC4340; MUID:943334383; PMID:8056840
;Accession: PC4341
;Molecule type: mRNA
;Residues: 1831-2029 <ASA>
;Genetic code: SGC5
;Introns: 57/3
;Superfamily: dynein heavy chain, cytosolic

Query Match      2.1%; Score 124; DB 2; Length 4540;
Best Local Similarity 18.5%; Pred. No. 10;
Matches 127; Conservative 111; Mismatches 214; Indels 234; Gaps 33;

Y 352 LSNLPNLTGARRLVEIIFLGRSPTS-----GPLCRTHRLSRRY 391
b 356 LSNIGQTYIPRLS-QLAESRSRELTNEMIKYFGQILHIKYDFKGLIKITQEPSQW 414
Y 392 WQWRPLFOQLLYN---HACQYVRLLRSHCRFRYANQQVTDALNTSPPLMDLLRLHSSP 448
b 415 DEEYKIFKQSIKVSQVHVKQYQGF--SHIKLOXQIQ-----HIQRLREHNEH- 460
Y 449 WQVYGLRACLCKVYSASLWGRHNERP-----FKNLKFIKGLKYGKLSLQ 496
b 461 -----LKEVIEQIIQNDQOEKENVQFATQETQAYDIFKNVEVF-DLSRDGDEQFF 513
Y 497 ELM--WKMKVECHW-----LRSSPGKQVFAAHRERILATF--LF-----WIMDT 541
b 514 RALKQVEIAIESVEATITNLRDLSGS---ASSAKEMFRILAKFNKLFSPRIKGAIQE 569
Y 542 YVQLLRSPFYITES-----TPQK---NRLFYVRK-----SVMSKLSQISQVROHLE 584
b 570 YQSOLLKTVHDKIQSLQNKFKETQYKQNSRLASARDIPLTSGFVIVSKQLQIRLQKYM 629
Y 585 RV-----LRLELQ-----BEVRHQDTWLAMPICRLRF 613
b 630 KVEQILGPQWAEITGKKCKENGTEFERILDSPGPALEDWKEINHN-----KA 678
Y 614 IPKNGLRPIVNMVSMGTRALGRKQKQHTQRLKTLFSLMANYERTKGPLHMGSSVLGM 673
b 679 VSQNEKLFVWTRRRGLRIVRYEKKLSQLF-KEVRNLSNM-----KTKVPYS-SHIA-- 730
Y 674 NDYRTWRALFVLVRALDTPRYFVKADYTCAYDAIPQGLVEVWAMHISESTYCIIR 733
b 731 NDAKASY-PFAL---SLQESLHYI---CITSQUNA-KSAKLIV----- 765
Y 734 QXAVVRDQSGGVHKSFRQVTTLSDLQFYMGQFLKHL----- 771
b 766 --AALARKEVQLQIQGGFNLYLTHKTLQLPYVKKFTDKVFELEQAVNGLNERIGIESLCE 823
Y 772 -----QDSALRNSVVIOSISWNSSSLPFDFFLHFLRHSVVKIGRCYTQCOGIP 825
b 824 AMKTCVPVDSLADKLDIQIEDVSLCFNFSN-----LHIWIQIDIX----- 864
Y 826 QGSSSLTLLCSLCFGDMENKL-----FAEVQRDGLL-----LRFVDDFLLVTPHLDQA 873
b 865 ---QIESILCDRTVQMKELWLNQFINYQKIQERGLVNVTVVHELKLQDIQIYVDPVEYA 921
Y 874 KTLSTLVHGVPEYGMINLQKTVN 899
```

```
Db 922 KYWFQEPHOM--IGQICSLPRLVAN 945

RESULT 13
E71436
hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002
C:Accession: E71436
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D.
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G.
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomene
erholt, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis th
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: E71436
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-2088 <BEV>
A:Cross-references: GB:297342; NID:G2245031; PID:G2245044
C:Genetics:
A:Map position: 4COP9-4G3845

Query Match      2.1%; Score 121; DB 2; Length 2088;
Best Local Similarity 19.9%; Pred. No. 5.3;
Matches 151; Conservative 82; Mismatches 251; Indels 276; Gaps 37;

QY 88 LCERNERNVLAQFELLNEA---RGPPMAFTSSVRSYLPNTV- 127
Db 1293 LSEKRRHIVETGLTLLTQASVPREYWPYAFAAV--YLINRMPTPLVSMESPFQKLFPS 1350
QY 128 ---TETLRVSGA---WM-----LLSRVGDLLVLLAHALY- 159
Db 1351 KPNYERLURVFCGLCPWLRPYTHNKLEERSRCVFCYLSLTQTAYLCFVHEKRLYTSRH 1410
QY 160 ---LLVP-----PSCAYQVCGSPLY-QICATTDIWPVSASVYRTPRPGRNPT 203
Db 1411 VVPDEAGFPFSSNLTSQNSLPTVTFEQSSPLVTPILSSSSVLPSCLS--PCTVLHQOQP 1468
QY 204 NLRFLQIKSSRQEAEPKPLALPSRGTKRLSLTSTSVPS-----A 244
Db 1469 PVTTPNSFHSSQPTTSAPLS-PHRTTMDTFQVQPTAPENENGPEPEAOSPPIGLSNPT 1527
QY 245 KKARCYVPVRVEEGPHRQVLPSPGKSWVPSPARSPEVPTABKDLSSKGVSDLSLSGV 304
Db 1528 HEATIGLPNPNRNPTEIEPT-----PAPHPKVPKTTTTTTTNRRTVSDAS----- 1575
QY 305 CCKHKPSTSLSPRONAFQLPFPFTRHFLYSRGDQO-ERLNPFSILLSNLPNLTGAR 363
Db 1576 ---HQPTA-----POON-----QENMKTRAKNNIKKPNTKFSLTATLPNRS--- 1613
QY 364 RLVEIIFLGSRPRTSGPLCRTHLSRRYW-----MRPLFQQLLVNHA 406
Db 1614 -----PSEPTNTVQALKKKWRPAMSDDEFDAQQRNHTWDIVPHESQLLVG--- 1658
QY 407 ECQVRL-----RSHCRFTANQO---VTDALNTSPHLLMDLRL-----HSSP 448
Db 1659 -CKWVFKLVLPNGAIDKYKARLVAKGFNQGVDAETFPVVKSTTIRLVLVDVAVKDD 1717
QY 449 WQVYGLRACLCKVYSASLWGRHNERPFFNKLKFIKGLKYGKLSLQELMWKMKVEDCH 508
Db 1718 WEIKQL-----DVNNAFLOGTITE- 1738
QY 509 WLRSSPG---KDRVPAAEHRLRERILATFLFLMDTYVQVLLRSFFYI-TESTFQKRLF 564
Db 1739 YMAQPPGFIDKOR-PTHVCELRKAIYG--LKQAPRAWYMEKQHLFNIGFVNSLSASLIF 1795
QY 565 PYKSVSKQSQISGVROHLRERLRELSQBEVHHQDTWLAMPICRLRIPKENGRLPIV 624
```

```
> 1796 IY-----WSDKSSID-----AVLTSIAE-----RF-----SIKDEP 1821
> 625 NMSVSMGTRALGRKQKQHTQR--LKTLPMSLNYERTK-----HPHLMGSSVLGNM 674
> 1822 DLHYFLGIEAT-RTKQGLHLMQRYKIDLLAKHNMAADAKPVLTPLTSPKLTUHGKTILN 1880
> 675 DI--YRTWRAPVLVRALDQTPRMVFVKADVTGAYDAIPQ 712
> 1881 DASYRS-----VGSQLYLATRPDIAYAVNLSQ 1911
>
> RESULT 14
> 10740
> Species: Caenorhabditis elegans
> Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
> Accession: T20740; T23800
> Submitted to the EMBL Data Library, December 1995
> Reference number: Z19317
> Accession: T20740
> Status: Preliminary; translated from GB/EMBL/DBSJ
> Molecule type: DNA
> Residues: 1-1646 <WIL>
> Cross-references: EMBL:Z68297; PIDN:CAA92597.1; GSPDB:GN00022; CESP:Flila10.4
> Experimental source: clone Flila10
> Steward, C.
> Submitted to the EMBL Data Library, January 1996
> Reference number: Z19800
> Accession: T23800
> Status: Preliminary; translated from GB/EMBL/DBSJ
> Molecule type: DNA
> Residues: 1-1646 <W12>
> Cross-references: EMBL:Z68507; PIDN:CAA92830.1; GSPDB:GN00022; CESP:Flila10.4
> Experimental source: clone M18
> Genetics:
> Gene: CESP:Flila10.4
> Map position: 4
> Introns: 102/3; 146/3; 265/3; 343/3; 424/1; 460/2; 736/2; 796/3; 891/2; 947/3; 1331/3;
>
> Query Match 2.08; Score 120.5; DB 2; Length 1646;
> Best Local Similarity 19.38; Pred. No. 4.1;
> Matches 244; Conservative 159; Mismatches 412; Indels 449; Gaps 64;
>
> 49 TLVAQCLVCMHWSGQPP-----PADLS-----PHQVS-----75
> 410 TVQQCLVCTFFSSDQENEDRQEDGGPGFLSGLWVPVEHLTSKTLILDSLRMDA 469
> 76 -----SLKELYARVVQRLC-----ERNERNVLAFGPELLNEARGGPPMAF 115
> 470 VAIEGVLRCCVACDMTQAVTAIDKLCVLVDENSEAGSBETKLEIAKAVANSQPSI 529
> 116 TSSVRSYLPNTVITLRSVGAWMLLLS-----RVGDDL-----LVYLAHCAI-----YL 160
> 530 LAAIGSLAASDDEI--VSPOLLCLSTLISACRGVADADLHRSVYVLAIVSLSPSYL 587
> 161 -----LVPPSCAYQVCGSPLYQCATDTHPS---VSASYRTPRPVGRNFTNLRFLQIK 212
> 588 NQFAGIAPPSPANK-RDAPVSEQVLEAMPSTQVTAAGPP-----628
> 213 SSSROEAPKPLALPSRGTKHLSLTSTSVPSAKARCYVPVVEEGPHRQVL--PTPSGK 270
> 629 -----CPCPVSTELWNKQVL-LTSKNLQAARTFIATITIKELNNLWYLCMATCEHL 681
> 271 SWVSPARSPEVTPAEKD-----LSSKGKVSDSLSSGVCCKKXPSSTLSLSPR 320
> 682 SWLLA-WRPQVQOFERETRDHSHNPTVVVNNALSDIAMLSLMDKVAPAIAALPD---737
> 321 QNAQLAPFIEIETHFLYSRGD-----QGE-RLNP-----SFLLSNLOPLNLTGAR 363
> 738 -----KSFLUSVIDALIRLDESILVAATGRESLPLAVLYKVCVLSLMRLQ-----784
> 364 RLVEIIFLGRSPRTSGPLCTRHLRRYV-----QMRPLFQQ 400
```

```
Db 785 -----VFVQKAAHHFKVCNHTSVSMRDWAVALVTSIAKHAVKSKTSMDAKSQOEMIIS 839
Qy 401 LL-----VNHAEQCVYVLLRSHCRFTANQVTDALNTSPPHLMDL-----441
Db 840 LLALCSIPHIQ---VRRQLDCVMSLMQTDGAFLLSTSWPNVIQIISAIIDSDTECELSL 896
Qy 442 -----LRLHSSPHQVYGFLEA-----CLCKVVSASLWGTNRNRRFFKMLKKFISLGKY 490
Db 897 VRQGYLGLRLVSS-----DFLQSIFFDCISGLVERAI---SRYSKONTQONI-----939
Qy 491 GKLSLQELMWKM-----KVEDCHWLRSRPGKDRVPAAEHLRLRERILATFLFLWMDTYVV 544
Db 940 -SLSALTLLWTISDFYRKYE-----SVGND-----ASEAVMVV-----972
Qy 545 QLLRSFYITESTFOKNRLFFYRKSV-WSKQSGVROHLERVLRLBELSQEVRHQDTW 603
Db 973 ---LVTCLSESCVDSR--FAVRKSAQCOTLLQTVTAHGHALR-----SAAWHSVI 1017
Qy 604 -LAMP1-----CRLRFIPKPNGLRPIVNNMSYNGTRALGRRKQAOHFTQRLKTLFSLM 655
Db 1018 QIMIPLLDKVRSQTRCASTEKNG---ELIMHSDTE---QKQWETETCIHTISAIKIP 1071
Qy 656 NYERTKHPLMGSSVLGMNDIYRTWRAPV--LRVRALDQTPRMVYFKADVTGAYDAIPQ 713
Db 1072 NSOR-----KSLIALNDFGAVWEAFGLVLDAAACVYENAEI---SLSAIRSQYEVLLG 1120
Qy 714 KLVEVVANNIRHSESTYCTIRQYAVVRDSQGVHKSFRQVTTLSDLQPYMGQ-----766
Db 1121 KISSQTLNVNSHEKN-----GSDS-----TIDAITPELPQAOQWESW 1158
Qy 767 -----FLKHLQSDSALRNSVYIE-QSIS-----MNESSSSLPDPFFLFLRHSVVKIGDR 816
Db 1159 KWLRIISRGLARQGAAMANSYNADSKSISSTPRMNSSSSL-----ASLAPG 1206
Qy 817 CYTQCOGIPQGSLSLTLLCSLCFGDMENKLPFAEVORDGILLRFVDDFLVTPHLDQAKTF 876
Db 1207 IY-----VFGPSHLTAIL--HVFPFLFDKVAKSIT-----VDDL-----1238
Qy 877 LSTLVHGVPEYGCMLNOKTVNVNPFVPGTLGGAPYOLPA---HCLFPWGLLLDQTQL 933
Db 1239 -----KYESLPAVLESMMNVIP---SEQAPFVLPSATHTLPTQALLAEAVKIV 1285
Qy 934 EYFCDSYGAQTSIKTSLTFQSVFKAGTKMKNKLLSVLRKCHGLFDLQVNSLQTVGIN 993
Db 1286 FVECTLSG---TSLRAAIPDQ-----IRLLKLFASMATOIRIS 1319
Qy 994 IYKIFL--LQAYRFHACVQLPFDQ-RVRKNLTFLGLIISQO---ASCCVAILKVKNPG 1046
Db 1320 PNKAPGGQKSTRDALTIVFSEYSIRIAIEFTTS--TSQNEVANSLIAIDIIKFLG 1377
Qy 1047 MTLKASGSPFPPAAHWCYQAFLLKLAHSHVYKCLGLPRLTAOKLLCRKLPEATMTILK 1106
Db 1378 -----EPLYMKYTCISASTWKLAAITS-----LMSVLRITSIP-YARQNPFVEFRGLWS 1422
Qy 1107 AAAD 1110
Db 1423 ATCD 1426
```

## RESULT 15

```
Tl1216
reverse transcriptase homolog ORF544 - red alga (Porphyra purpurea) mitochondrion
C/Species: Mitochondrion Porphyra purpurea
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C/Accession: Tl1216
R/Burger, G.; Saint-Louis, D.; Gray, M.W.; Lang, B.F.
submitted to the EMBL Data Library, December 1998
A/Description: Complete sequence of the mitochondrial DNA of the red alga, Porphyra pu
A/Reference number: Z17255
A/Accession: Tl1216
A/Status: preliminary; translated from GB/EMBL/DBSJ
A/Molecule type: DNA
```



```
;Residues: 1-544 <BUR>
;Cross-References: EMBL:AF114794; NID:G4106927; PID:G4106928; PIDN:AAD03095.1
;Genetics:
;Genome: mitochondrion
;Note: orf544
;Keywords: mitochondrion

Query Match      2.0%; Score 116.5; DB 2; Length 544;
Best Local Similarity 21.7%; Pred. No. 1.6;
Matches 75; Conservative 41; Mismatches 94; Indels 135; Gaps 16;

Y 573 KLGSGVGRQHLEFVRLRELSOBEVRRHODTWLAMPICRLRFIPKPNGLRPIVNMYSMGT 632
b 70 KISSITPDERMELVRNIHVQNSDK-----ILRVT-IPKPNG-----SV 107

Y 633 RALGRKQAQHTQRLKTLFSLNLYERTKHPHLMGSSVLGMNDIYRTWR-----AFVLRVR 688
b 108 RNLGIPTIRDRAKQCLVKFALEPQVEAIFEPNSYG-----FRPGRSANDARKAIVK 158

Y 689 ALDQTPMYFVKADVTGAYDAIPQGLVVEVYANNIRHSESTYCIROYAVVRDSQGVHK 748
b 159 CLQRPK-HILDADINGCFDNIHDKLEKL-----NIFPL----- 193

Y 749 SPFRQVTT-----LSDLQPYMGQFLKHLQSDASALENSVVIQISMSNESSSLFDFP 802
b 194 -FRNQINTWLKVGVLSDFOG-----NKTETIIPQS----- 221

Y 803 LHFLRHSVVKIGRCYTCQCGIPOGSSSLTLCSLCFGDMENKLFQVQSDGLLREVD 862
b 222 -----GTPQGGTISPLLANIALHGIE-KLVS--KGVYLIRYADD 258

Y 863 FLIV---TEHLDQAK-----TFLSTLVHGVPYEGCMINLQKTWNF 900
b 259 FLVLCNEEREKELLEAKQKIEVFLKTI-----GLEFSAEKTIKITY 296
```

Search completed: February 4, 2004, 14:12:28  
Job time : 32 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

(protein - protein search, using sw model)

on on: February 4, 2004, 14:09:14 ; Search time 18 Seconds  
(without alignments)  
2931.331 Million cell updates/sec

tle: US-09-042-460-2

fect score: 5901

quence: 1 MTRAPRCPAVRLRSRYE.....TILKAADPALSTDFQITLD 1122

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 127863 seqs, 47026705 residues

tal number of hits satisfying chosen parameters: 127863

nimum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase : SwissProt.41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

slut	No.	Score	Query	Length	DB	ID	Description
1	5901	100.0	1122	1	TERT_MOUSE	070372	mus musculus
2	3505	59.4	1132	1	TERT_HUMAN	014746	homo sapien
3	621.5	10.5	988	1	TERT_SCHPO	013339	schizosacch
4	426.5	7.2	1132	1	TERT_OXYTR	076332	oxytricha t
5	410.5	7.0	1031	1	TERT_EUPAE	009939	euploetes ae
6	401.5	6.8	1117	1	TERT_TETTH	077448	tetrahymena
7	341	5.8	884	1	TERT_YEAST	06163	saccharomyc
8	131	2.2	778	1	TAST_HUMAN	012815	homo sapien
9	124	2.1	4540	1	DYHC_PASTE	021717	paramacium
10	119	2.0	873	1	GLND_VIBU	08db93	vibrio vuln
11	116	2.0	881	1	GLND_VIBCH	09xpvo	vibrio chol
12	116	2.0	968	1	SNXD_HUMAN	09y5w8	homo sapien
13	114.5	1.9	1199	1	Y173_HUMAN	014679	homo sapien
14	113.5	1.9	653	1	DEND_RAT	P50617	rattus norv
15	111.5	1.9	1415	1	ICPA_HSMVG	002362	marek's dis
16	110	1.9	812	1	EBN2_EBV	P12977	epstein-bar
17	108	1.8	880	1	ARG_YEAST	P05085	saccharomyc
18	108	1.8	983	1	4ET_MOUSE	Q9est3	mus musculus
19	106	1.8	958	1	MSH4_MOUSE	Q99mt2	mus musculus
20	105.5	1.8	901	1	R8L3_ARATH	Q9fjb5	arabidopsis
21	105	1.8	1473	1	NALI_HUMAN	Q9c000	homo sapien
22	105	1.8	1493	1	M3K1_RAT	Q62925	rattus norv
23	104.5	1.8	916	1	RTUK_DROFU	P21329	drosophila
24	104.5	1.8	3511	1	MYL5_MOUSE	Q9qz44	mus musculus
25	103	1.7	854	1	YC19_HUMAN	Q9ulk2	homo sapien
26	102.5	1.7	788	1	CY14_NEUCR	P23622	neurospora
27	102.5	1.7	2335	1	TORI_SCHPO	014356	schizosacch
28	102.5	1.7	2482	1	VWF_FIG	Q28833	sus scrofa
29	102	1.7	1493	1	M3K1_MOUSE	P53349	mus musculus
30	102	1.7	1795	1	ESPI_HUMAN	Q14674	homo sapien
31	102	1.7	1815	1	SHK3_RAT	Q9jlua	rattus norv
32	101	1.7	677	1	T2D5_HUMAN	P49848	homo sapien
33	101	1.7	994	1	MERK_MOUSE	Q60805	mus musculus

34 101 1.7 3530 1 MY15\_HUMAN  
35 100.5 1.7 1214 1 SIPI1\_HUMAN  
36 100 1.7 846 1 AMDM\_SCHPO  
37 100 1.7 2280 1 YCF2\_OENHO  
38 99.5 1.7 759 1 AL16\_SCHPO  
39 99.5 1.7 810 1 AMDM\_YEAST  
40 99.5 1.7 867 1 PML1\_MOUSE  
41 99.5 1.7 5596 1 MN1\_HUMAN  
42 98.5 1.7 519 1 TRPE\_BUCDN  
43 98.5 1.7 1369 1 MSH5\_CAEEL  
44 98.5 1.7 2163 1 BRR2\_YEAST  
45 98 1.7 745 1 CUL2\_HUMAN

#### ALIGNMENTS

RESULT 1  
TERT\_MOUSE STANDARD; PRT; 1122 AA.  
AC 070372; 035432;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic subunit).  
GN TERT.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98241176; PubMed=9582020;  
RA Greenberg R.A., Allsopp R.C., Chin L., Morin G.B., Depinho R.A.;  
RT "Expression of mouse telomerase reverse transcriptase during development, differentiation and proliferation."  
RL Oncogene 16:1723-1730(1998)  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98393668; PubMed=9724727;  
RA Martin-Rivera L., Herrera E., Albar J.P., Blasco M.A.;  
RT "Expression of mouse telomerase catalytic subunit in embryos and adult tissues."  
RL Proc. Natl. Acad. Sci. U.S.A. 95:10471-10476(1998).  
RN [3]  
RP SEQUENCE OF 550-616 FROM N.A.  
RA Drissi R., Cleveland J.L.;  
RT "Partial sequence of Mus musculus telomerase catalytic subunit homolog."  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT LONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS SINGLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
CC -!- SUBUNIT: Interacts with PINK1 (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.  
CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AF051911; AAC09323.1; -  
CC EMBL; AF073311; AAC34821.1; -  
CC EMBL; AF029235; AAB84200.1; -  
CC MGD; MGI:1202709; TERT.  
-----

R InterPro; IPR000477; RVTse.  
R InterPro; IPR003545; Telomerase\_RT.  
R Pfam; PF00078; rvt; 1.  
R PRINTS; PR01365; TELOMERASERT.  
W Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
T DNA-binding. 553 553 I -> V (IN REF. 3).  
Q CONFLICT 1122 AA; 127977 MW; F85266905DD6558C CRC64;  
SEQUENCE

Query Match 100.0%; Score 5901; DB 1; Length 1122;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 MTRAPCPAVRSLLSRVREVPWPLATFVRRIGPEGRRVQGDPKIYRTLVVAQCLVCMHW 60  
b 1 MTRAPCPAVRSLLSRVREVPWPLATFVRRIGPEGRRVQGDPKIYRTLVVAQCLVCMHW 60

Y 61 GSOPPPADLSFHQVSLKELVARVQVRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120  
b 61 GSOPPPADLSFHQVSLKELVARVQVRLCERNERNVLAFFGELLNEARGGPPMAFTSSVR 120

Y 121 SYLPNTVITLTVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPYQICA 180  
b 121 SYLPNTVITLTVSGAWMLLSRVGDDLLVYLLAHCALYLLVPPSCAYQVCGSPYQICA 180

Y 181 TTIDWPSVASRYPTRPVGRNTNRLFLQIKSSROEAPKPLALPSRGTKRHLSLTSTS 240  
b 181 TTIDWPSVASRYPTRPVGRNTNRLFLQIKSSROEAPKPLALPSRGTKRHLSLTSTS 240

Y 241 VPSAKKARCPVPRVEEGPHRQVLTTPSGKSWPSPARSPEVPTAEKDLGKGVDSLSL 300  
b 241 VPSAKKARCPVPRVEEGPHRQVLTTPSGKSWPSPARSPEVPTAEKDLGKGVDSLSL 300

Y 301 SGSVCKKHPSTSLSPRONAFOLRPIETRHFLYSKGQGERLNPFSLLNLQNLTL 360  
b 301 SGSVCKKHPSTSLSPRONAFOLRPIETRHFLYSKGQGERLNPFSLLNLQNLTL 360

Y 361 GARLVEIIFLGSRRPTSGFLCRTHLSRRYQWMPFLFQQLLVNHAECQVYVLLRSHCRF 420  
b 361 GARLVEIIFLGSRRPTSGFLCRTHLSRRYQWMPFLFQQLLVNHAECQVYVLLRSHCRF 420

Y 421 RTANQVTTALNTSPHLMDLRLHSSPMQVYGFRLACLVKVSASLWGRHNRERFFKN 480  
b 421 RTANQVTTALNTSPHLMDLRLHSSPMQVYGFRLACLVKVSASLWGRHNRERFFKN 480

Y 481 LKXFISLGYKLSLQELMKNKVEDCHWLRSPPKDRVPAAEHLRLERILATFLWLMD 540  
b 481 LKXFISLGYKLSLQELMKNKVEDCHWLRSPPKDRVPAAEHLRLERILATFLWLMD 540

Y 541 TYVQLLRFFYITESTFQKRLFFYRKSVWSKLOSIGVROHLERVLRELQSEVRHHQ 600  
b 541 TYVQLLRFFYITESTFQKRLFFYRKSVWSKLOSIGVROHLERVLRELQSEVRHHQ 600

Y 601 DTWLAMPICRLRPIKPNGLRPIVNMVSMGTRALGRKQKQHTQRLKTLFSLMYERT 660  
b 601 DTWLAMPICRLRPIKPNGLRPIVNMVSMGTRALGRKQKQHTQRLKTLFSLMYERT 660

Y 661 KHPLMGSSVLGMNDIYRTWRFAVLVRALDTPRMVFKADVTDGAYDAIPQGLVEVVA 720  
b 661 KHPLMGSSVLGMNDIYRTWRFAVLVRALDTPRMVFKADVTDGAYDAIPQGLVEVVA 720

Y 721 NMIRHSESTYCIQAVARDSDGQVHKSFRRQVTTLSLOPYMGQFLKHLQSDASALR 780  
b 721 NMIRHSESTYCIQAVARDSDGQVHKSFRRQVTTLSLOPYMGQFLKHLQSDASALR 780

Y 781 NSVVFQSIISMNESSSLDFLHLRHSVVKIGRCYTCQCGI:PGQSSLTLLCSLCFG 840  
b 781 NSVVFQSIISMNESSSLDFLHLRHSVVKIGRCYTCQCGI:PGQSSLTLLCSLCFG 840

Y 841 DMENKLPFAEVQDGLLLRVDDFLVTHLDOAKTFLSLVHGVPYEGCMINLQKTVNF 900  
b 841 DMENKLPFAEVQDGLLLRVDDFLVTHLDOAKTFLSLVHGVPYEGCMINLQKTVNF 900

QY 901 PVEPGLTGGAAAPYQOLPAHCLFPMWGLLLDTQTLVEFCDYSGYAQISIKTSLTFQSVFKAG 960  
Db 901 PVEPGLTGGAAAPYQOLPAHCLFPMWGLLLDTQTLVEFCDYSGYAQISIKTSLTFQSVFKAG 960

QY 961 KTMENKLLSVLRKCHGLFLLDQVNSLQTVNCINIKIFLLQAVRPHACVQLPFFDORVRK 1020  
Db 961 KTMENKLLSVLRKCHGLFLLDQVNSLQTVNCINIKIFLLQAVRPHACVQLPFFDORVRK 1020

QY 1021 NLTPFLGIISQASCCYAILKVNKPGMTLKASGSPPEAAHMLCYQAFLLKLAHSHVIYK 1080  
Db 1021 NLTPFLGIISQASCCYAILKVNKPGMTLKASGSPPEAAHMLCYQAFLLKLAHSHVIYK 1080

QY 1081 CLLGPLRTAQKLLCRKLPEATWILKAADPALSTDFQILD 1122  
Db 1081 CLLGPLRTAQKLLCRKLPEATWILKAADPALSTDFQILD 1122

RESULT 2  
TERT HUMAN  
ID TERT HUMAN STANDARD; PRT; 1132 AA.  
AC 014746; O14783;  
DT 30-MAY-2000 (rel. 39, Created)  
DT 30-MAY-2000 (rel. 39, Last sequence update)  
DT 28-FEB-2003 (rel. 41, Last annotation update)  
DE Telomerase reverse transcriptase (EC 2.7.7.-) (telomerase catalytic subunit) (HEST2).  
DE subunit (HEST2).  
GN TERT OR TERT OR EST2 OR TCSI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=97400623; PubMed=9252327;  
RA Nakamura T.M., Morin G.S., Chapman K.B., Weinrich S.L., Andrews W.H., Lingner J., Harley C.B., Cech T.R.;  
RT "Telomerase catalytic subunit homologs from fission yeast and human.";  
RL Science 277:955-959(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97433088; PubMed=9289757;  
RA Meyerson M., Counter C.M., Eaton C.N., Ellisen L.W., Steiner P., Caddle S.D., Ziaugra L., Beijersbergen R.L., Davidoff M.J., Liu Q., Bacchetti S., Haber D.A., Weinberg R.A.;  
RT "hEST2, the putative human telomerase catalytic subunit gene, is up-regulated in tumor cells and during immortalization.";  
RL Cell 90:785-795(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97433088; PubMed=9289757;  
RA Wick M., Zubov D., Hagen G.;  
RT "Genomic organization and promoter characterization of the gene encoding the human telomerase reverse transcriptase (hTERT).";  
RL Gene 232:97-106(1999).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Londono-Vallejo J.A.;  
RT "Sequence of a BAC carrying the entire hTERT gene.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
CC -!- SUBUNIT: Interacts with PINK1.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- DISEASE: ACTIVATION AND CANCER CELL PATHOGENESIS. IMMORTALIZATION OF TELOMERASE HAS BEEN IMPLICATED IN CELL IMMORTALIZATION AND CANCER CELL PATHOGENESIS.  
CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY. TELOMERASE SUBFAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF015950; AAC51672.1; -  
EMBL; AF018167; AAC51724.1; -  
EMBL; AF128894; AAD30037.1; -  
EMBL; AF128893; AAD30037.1; JOINED.  
EMBL; AY007685; AAG23289.1; -  
PIR; TC3844; TC3844.  
Genew; HGNC:11730; TERT.  
MIM; 167270; -  
GO; GO:0005696; C:telomere; TAS.  
GO; GO:0003721; F:telomeric template RNA reverse transcriptase. . .; TAS.  
GO; GO:007003; P:telomere binding; TAS.  
InterPro; IPR000477; RVTse.  
InterPro; IPR003545; Telomerase\_RT.  
Pfam; PF00078; tvt; 1.  
PRINTS; PR01365; TELOMERASERT  
Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
DNA-binding. 516 516 D -> G (IN REF. 2).  
CONFLICT 1132 AA; 126996 MW; 94E35469C4CA33A0 CRC64;  
SEQUENCE

Query Match 59.4%; Score 3505; DB 1; Length 1132;  
Best Local Similarity 62.4%; Pred. No. 2.5e-251;  
Matches 719; Conservative 122; Mismatches 260; Indels 52; Gaps 13;

1 MTRAPCAVSLRLSYREVMPLATVRLGPEGRLVQDPKLYRTLVAQLVCMHW 60  
2 1 MPRAPCRVSLRLSYREVLPLATVRLGPEGRLVQDPKLYRTLVAQLVCMHW 60  
3  
4 61 GSOPPADLSFHVSSKELVARVQVRLCERNENLVAFGLFELLNEARGGPMWFTSSVR 120  
5  
6 61 DARPPAAPSPRQVSKELVARVQVRLCERNENLVAFGLFELLNEARGGPMWFTSSVR 120  
7  
8 121 SYLPTNTVTLRVSGAMMLLSRYGDDLLVYLLAHCALYLLVPFSCAYQVCGSPLYQICA 180  
9 121 SYLPTNTVTLRVSGAMMLLSRYGDDLLVYLLAHCALYLLVPFSCAYQVCGSPLYQICA 180  
10  
11 181 TTDWPSVASYRTPVGRNFTNRLFQIKSSROEAPKPLALPSRGTKRLSLTSTS 240  
12 181 ATQARPPPHAS-GPRRLG-----CERAWNHSVEAGVPLGLPAPGARRGGASRS 231  
13  
14 241 VPSAKKACVPPVVEEGP-----HRQVLPSPGKSM-VPSFAPSPEVPTAEKDLSSK 292  
15 241 VPSAKKACVPPVVEEGP-----HRQVLPSPGKSM-VPSFAPSPEVPTAEKDLSSK 292  
16 232 LPLKPRGGAAPERTPVCGSWAHPGRTGPDGFCVSPAR-----PAEETSLE 286  
17 232 LPLKPRGGAAPERTPVCGSWAHPGRTGPDGFCVSPAR-----PAEETSLE 286  
18 293 GKVDLSLS-GSVCKKHPSTSLSPRNAPQLRP-FIETHFLYSRGDGOERLNPSF 350  
19 293 GKVDLSLS-GSVCKKHPSTSLSPRNAPQLRP-FIETHFLYSRGDGOERLNPSF 350  
20 287 GALSCTRHSFVSQRHAGPSTSRPPEPMDTPCPVVAETKHFYLSGD-KEQLRPSF 345  
21 287 GALSCTRHSFVSQRHAGPSTSRPPEPMDTPCPVVAETKHFYLSGD-KEQLRPSF 345  
22 351 LLSLNQPNLTGARLVEITFLGSRPSTSGPLCRHLSRYWMBPLFOQLVNVHAEQY 410  
23 351 LLSLNQPNLTGARLVEITFLGSRPSTSGPLCRHLSRYWMBPLFOQLVNVHAEQY 410  
24 346 LLSLRPSLTGARLVEITFLGSRPSTSGPLCRHLSRYWMBPLFOQLVNVHAEQY 405  
25 346 LLSLRPSLTGARLVEITFLGSRPSTSGPLCRHLSRYWMBPLFOQLVNVHAEQY 405  
26 411 VALLRSHCRFTANQVQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
27 411 VALLRSHCRFTANQVQVTDAL-----NTSPHMLDLRLHSSPWQY 452  
28 406 GYLLKTHCPRAA---VTPAAGVCAKPKQGSVAAPPEEDTPRRLVQLLRQHSPPWQY 462  
29 406 GYLLKTHCPRAA---VTPAAGVCAKPKQGSVAAPPEEDTPRRLVQLLRQHSPPWQY 462  
30 453 GFLRACLCKVYSASLWGRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWLS 512  
31 453 GFLRACLCKVYSASLWGRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWLS 512  
32 463 GEVRACLRLVPPGLWGRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWLS 522  
33 463 GEVRACLRLVPPGLWGRHNERFFKNLKFISLGYKLSLOELMWKMKVEDCHWLS 522  
34 513 SPGRDVPAAEHLRERLIALFELWMDTVVQLLRSFFVITESTFOKRLFFPKYSWS 572  
35 513 SPGRDVPAAEHLRERLIALFELWMDTVVQLLRSFFVITESTFOKRLFFPKYSWS 572  
36 523 SPGVCCVPAEHLRERLIALFELWMDTVVQLLRSFFVITESTFOKRLFFPKYSWS 582  
37 523 SPGVCCVPAEHLRERLIALFELWMDTVVQLLRSFFVITESTFOKRLFFPKYSWS 582  
38 573 KQISGTVQHLERVLRLSELSEFVHHQDWTWAMPICRLRFPKPNGLRPIVMSYMT 632  
39 573 KQISGTVQHLERVLRLSELSEFVHHQDWTWAMPICRLRFPKPNGLRPIVMSYMT 632

Db 583 KLSIGIRCHLXVQLRELSAEVQRHREARPALITSRLRPFKPDGLRPIVNDYVYGA 642  
Qy 533 RALGRKQKQHFQRLKTLFSLNTERYKHPHMGSSVLGNDIYRTWRAPVLVRALDQ 692  
Db 643 RTRRREKRAERLTSRYKALFSLNTERYKHPHMGSSVLGNDIYRTWRAPVLVRALDQ 702  
Qy 693 TRPMYFKADVTGAYDAIPQKGLVEVANNMHSSTYCIROYAVVRDSDQGVHKSFR 752  
Db 703 PPFLYFKVDVTGAYDTIPQDRLETVIASIILK-PONTYCVRYAVVQKAAGHVRKAFKS 761  
Qy 753 QVTTLSLOPYMGQFLKHLQSDASALRNVVIEOSISNWSSSSLFDFFLHPLRHSVVK 812  
Db 762 HVSTLIDLOPYMGQFLKHLQSDASALRNVVIEOSISNWSSSSLFDFFLHPLRHSVVK 819  
Qy 813 IGRCVTCOGIPQGSLSLTLCSLCFCGDMENKLPABVQRDGLLRLLRVDLFTVPHLDQ 872  
Db 820 IRKSVVQCGIPQGSLSLTLCSLCFCGDMENKLPABVQRDGLLRLLRVDLFTVPHLDQ 879  
Qy 873 AKTFLSTLVHGVPEYGCMLNLOKTVVNFVEPTGLGGAAPYQLPAPHCLFPMCGLLDQT 932  
Db 880 AKTFLSTLVHGVPEYGCMLNLOKTVVNFVEPTGLGGAAPYQLPAPHCLFPMCGLLDQT 939  
Qy 933 LEVPCDYSYAGTSIKTSITFSQVFKAGTKWKNLLSVLRKCHGLFLDLQVNSLOTVC 992  
Db 940 LEVPCDYSYAGTSIKTSITFSQVFKAGTKWKNLLSVLRKCHGLFLDLQVNSLOTVC 999  
Qy 993 NIYKIFLLQAYRFAHACVIOQLPDRVRKNTLFFLGTIISSQASCCYAILKVKNPGMTLKA 1052  
Db 1000 NIYKIFLLQAYRFAHACVIOQLPDRVRKNTLFFLGTIISSQASCCYAILKVKNPGMTLKA 1059  
Qy 1053 GS---PPPEAAHMLCVQAFLLKLAHSVYKCLLPRLTAQKLLCRKLPKPEATMILKAAA 1109  
Db 1060 GAAGPLFSEAVQMLCHQAFLLKLAHSVYKCLLPRLTAQKLLCRKLPKPEATMILKAAA 1119  
Qy 1110 DPALSTDFOTILD 1122  
Db 1120 NPALPSDFOTILD 1132

## RESULT 3

TERT SCHPO  
ID TERT SCHPO STANDARD; PRT; 988 AA.  
AC 013339; 013338;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic subunit).  
GN TRT1 OR SPBC29A3.14C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=97400623; PubMed=9252327;  
RA Nakamura T.M., Morin G.B., Chapman K.B., Weinrich S.L., Andrews W.H., Lingner J., Harley C.B., Cech T.R.;  
RT "Telomerase catalytic subunit homologs from fission yeast and human.";  
RL Science 277:955-959(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis B., Feltwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris B., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E., Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B., Welter J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fritz C., Holzer E., Moesl D., Hilbert H., Borzym K., Langer I., Beck A., Leinhardt R., Reinhardt R., Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B., Goffeau A., Cadieu E., Dreanc S., Gloux S., Lelaure V., Mottier S., Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M., Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado I., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L., Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J., Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 T "The genome sequence of Schizosaccharomyces pombe.";  
 L Nature 415:871-880(2002).  
 C -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR  
 C THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT  
 C ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS  
 C SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE  
 C SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
 C -!- SUBCELLULAR LOCATION: Nuclear.  
 C -!- ALTERNATIVE PRODUCTS:  
 C Event=Alternative splicing; Named isoforms=2;  
 C Name=1;  
 C IsoId=O13339-1; Sequence=Displayed;  
 C Name=2;  
 C IsoId=O13339-2; Sequence=VSP\_006395;  
 C -!- MISCELLANEOUS: DELETION CAUSES TELOMERE SHORTENING AND SENESCENCE.  
 C -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.  
 C TELOMERASE SUBFAMILY.  
 C This SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C  
 C EMBL; AF015783; AAC49802.1; -  
 C EMBL; AF015783; AAC49803.1; -  
 C EMBL; AL022299; CAAL8391.1; -  
 C PIR; T03838; T03838.  
 C GeneDB SPombe; SPBC29A3.14c; -  
 C InterPro; IPR000477; RVase.  
 C InterPro; IPR003545; Telomerase\_RT.  
 C Pfam; PF00078; rvt; 2.  
 C PRINTS; PR01365; TELOMERASERT.  
 C Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
 C DNA-binding; Alternative splicing.  
 C VARSP/LIC 524 524 K -> KQ (in isoform 2).  
 C /FTid=VSP\_006395.  
 C  
 C SEQUENCE 988 AA; 116328 MW; ABZDC7030228F443 CRC64;  
 C  
 C Query Match 10.5%; Score 621.5; DB 1; Length 988;  
 C Best Local Similarity 23.6%; Pred. No. 7.1e-38;  
 C Matches 253; Conservative 193; Mismatches 384; Indels 243; Gaps 45;  
 C  
 C 76 SLKELVARVQVLCERN---ERNVLAFFGELLNE-ARGPPMAFTSSVRSYLPNTVIETL 131  
 C 83 SQBELIANVVKQMFDSERRNLLMKGSFMHEDFRAMHNGVQNDLVSTFFNYLSIL 142  
 C 132 RVSGAWMLLSRVGDDLLVYLLAHALYLLVPFSCAYQVCGSPYQICATTDIMPSVAS 191  
 C 143 E-SKWQLLLEIGSDAMHYLLSKGSIFEALPDNDVQLSIGIFPK- 187  
 C 192 YRTRPVGRNFTNRLFLQIKSSRQEAEPKALPSRGTKRLHLSITSV-----PS 243  
 C 188 -----NNV--FEETVSKRKRTTETSITQNKARKEVSNWSISIRSFIVRSS 234

244 AKKARCVPVRVEGPHRQVLPTPSGKSNVPSPARSPVPTAEKDLSSKKGKYSILSLG-G 302  
 235 YKPK- 247  
 303 SVCKKXKFSSTLSLPPRQ---NAFLQAPPIETRHFLYSGDGOERLNPSPFLSLNLPN 358  
 248 SICDENTVHWMLQWIPRQGLINAFQVKQL 292  
 359 LTGARRLVEIFLGSRPRTSGPCRTHRLSRRWQVRPLFQQLLVNHAECQVRLLRSHC 418  
 293 ---KLLKVY---PLIEQTAKRLHRIS- 325  
 419 RFRANQOVTDALNTSPHMLDLRLHSSPQVYGFELACLCKVVSASLWGRHNRERF- 477  
 326 --DTHDE- 367  
 478 --FKNLKKFISLGGYKLSLOELMWKVEDCHWL-----RSSPGKORVPAAEHRLRERIL 531  
 368 IILKDLTEFLKLSRYESFLHYLMSNIKISEIEMVLGKRSNA---KMLSDPFEKQIF 424  
 532 ATFLFWMIDTVVQLLESFFVITESFQKRLFFYKSVMSKLSQSIGVCHLERVRL--- 588  
 425 AEFYWLNSFIPILOSFFYITESDOLNRVTYFRKDIWKL-----CRPFTSMKMEAF 480  
 589 RELSQEEVRRHQDTW-LAMPICRLRFPKNGLRPIVNMYSNGTALGRKROAOHFTOR 647  
 481 EKINENNVR--MDTQKTLPPAVIRLLPKNTFLITNLKRFLLKVGSKNKMVLSTNQT 538  
 648 LKTLFSLMYNERTKXPHLMG--SSVLGMN-DIYR---TWRAFVLRVALDQTPRYVFKVA 701  
 539 LRPVASILK-----HLINESGIPFNLEVTKLLTFKDLKLRHMPGR--KKYFVRI 589  
 702 DVTGAYDAIPQCKLVEVVMNIRHSESTYCIROYAVVRDSSQGVKHSFRQVTTLSLQ 761  
 590 DIKCYDRIKQDLMEFRVKKLKDPE--FVIRKYATIHATSD-RATKNFVSEAFSYDMV 646  
 762 PY--MGQFLKHLQSDASALNSVVIQSSIMNESSSLDFDLHLFLRHSVVKIGDRCYT 819  
 647 PFEKVQQLS-MKTS- -TLFVDFVDYWTYKSSSEIFKMLKEHLSGHIKIGNSQYL 699  
 820 CQCGIPQSSSLTLLCSLCFGDMENKLFABEVQDQ--LLLRFVDDFLVTPHLDQAFTLS 878  
 700 QKVGIPOQSILSSFLCHFYMEDLIDEVLSFKKKGSVLLVVDVDFLITVKKDAKFLN 759  
 879 TLVHGVPYEGCMINLQKTVNVPVPEOTLGAAPYQLPAHCLFPWCGLLDQTLEVP-- 936  
 760 LSLRGFEKHNFSTSEKTVINFENSNGIINN- -FFNESKKRMPPFFGFSYNMRLTLA 817  
 937 ---CDYSGYACTSIK-TSLTFQSVFKAGTKMRNKLVLRLKCHGLFDLDQVNSLQTVCI 992  
 818 CPKIDEALFNSTVELTKHMKGSFFY--KILRSLASFAQ-----VFIDITHNSKFNSSC 870  
 993 NIYKI-----FLQAYRFHACVIOLP-----FDQVRKNLTFPLGIITSSQASCC 1036  
 871 NIYRLGYSMCVRAQAYLKBKMDIFIPQRMFITDLLNVIGRKIKWKLAEILGYTSRFLSS 930  
 1037 YAILKVNQPGMT--LKASGSPFPAEAHWLCYQAFLLKLAHSVIYKCLLGLR 1087  
 931 AEVKWLFLGLMRDGLKPSFYHP-----CFEQLIYQFSLTDLIK-----PLR 973  
 RESULT 4  
 ID TERT\_OXYTR STANDARD; PRT; 1132 AA.  
 AC O76332;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic  
 DE subunit) (Telomerase subunit P133).  
 GN TERT.  
 OS Oxytricha trifallax.

543 VLKVFEDLAIITMRCYFYSTSKAKYORIFYYRKNINWMTIRUSIDDLKQ-NLKQVEK 601  
 594 EEVRHHODTFLWAMPICRLRFIPKPNGLRPINVMYSMTGTRALGRKQAQHTORLKTFLS 653  
 602 KEMRIFCESQNPAP-GKLRLIPKGDFTFPIMTFNKRKIPNQ-VGRFQSRWTTNNKLIQTAAH 659  
 654 ML-NYERTKPHLMGSSVLGMNDIYRTWRAFVLIRVRALDQTPRMVYKADVTGAYDAIPQ 712  
 660 MLRLKLSKMFKSGFGFAVFNDDIMKRYENFVQKWQIN-SPKLYFVAMDIKCYDNVD- 717  
 713 GKLVEVANNMIRHS---ESTYCTIQVAVVRDRSQGVHKS-FRRQVTTLSLQPYMGQFL 768  
 718 --CERVNVFLQKSLDMXDEYFILNTEVLKRNKNIIVERSNFRK---LPKQYFRYKFO 770  
 769 K-----HLQD--SDASALRNSVIEQISWNSSSSLPFFHLFLHSHSVVKI 813  
 771 KIGDGGSYPTLFEILDEDFNLN-MKRTIIVQEQRKFFPKNDLLOFVLKICNNYVTF 829  
 814 GDRCYTCQCIPOGSSLSLTLLCSLQCGDMENKLPFAVORD-----GLLIRFVDDFL 864  
 830 NKQYKQKKGIPGLCVSVILSYFYANLEENALQFLAKESMDPEKBEINLLMLTDDYL 889  
 865 LVTPHLDQAKTFLSTIVHGVPBYGCMINLQKTVNFPVPEPTLGGAAAPYQLPAHC----- 919  
 890 LMTTEKNAMLFIEKLYQLSLGNFFKFMKKLKTNFALNLQKIG-----CTNTTQ 939  
 920 -----LPPWGLLIDTQTELEVCDYS---GVAQTSIKTSLTFSQVFKAGKTWENKL 967  
 940 DIDINDLFWHIGISIDIKTINIQNIKEGILCTLVNMQTNESILWLKKUKLSFL 999  
 968 LSVLRXCHGLFLDLQVNSLQTVNCINIKFLLQAYRPHAC 1008  
 1000 MNNI-----SPYFKSTINTKQFANITLSKLVIATAAEKVAC 1035

RESULT 5  
 ID TERT EUPAE STANDARD; PRT; 1031 AA.  
 AC O00939;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Telomerase reverse transcriptase (EC 2.7.7.-) (Telomerase catalytic subunit) (Telomerase subunit P123).  
 DE Euplotes aediculatus.  
 OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;  
 OC Euplotida; Euplotidae; Euplotes.  
 ON NCBI\_TaxId=5940;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97274210; PubMed=9110970;  
 RX Lignier J., Hughes T.R., Shevchenko A., Mann M., Lundblad V.,  
 RA Cech T.R.;  
 RA "Reverse transcriptase motifs in the catalytic subunit of  
 RT telomerase".  
 RT Science 276:561-567(1997).  
 CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR  
 CC THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT  
 CC ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS  
 CC SINGLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE  
 CC SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.  
 CC TELOMERASE SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by www.isb-sib.ch/announcement/  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/  
 CC or send an email to license@isb-sib.ch).

DR EMBL; U95964; AAC47515.1; -.  
 DR InterPro; IPR000477; RVTse.  
 DR InterPro; IPR003545; Telomerase\_RT.  
 DR Pfam; PF00078; rvt; 1.  
 DR PRINTS; PR01365; TELOMERASERT.  
 CW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
 CW DNA-binding.  
 SQ SEQUENCE 1031 AA; 122562 MW; 57B87A63A1FED60F CRC64;

Query Match 7.0%; Score 410.5; DB 1; Length 1031;  
 Best Local Similarity 20.8%; Pred. No. 3.2e-22;  
 Matches 132; Conservative 128; Mismatches 265; Indels 109; Gaps 20;

2Y 470 TRINERFFKNLKKFTSLQYKGLSLOELMWKVKEDCHWLSSPGKDRVPAAEHRLER 529  
 376 TGRNRKFKVKKYVELANKHELHNKLLLEKINTRISMMQVTSAXHFFYDFHE-NIY 434  
 530 ILATFLFWMMDTVVOLLRSFFVITESTQKRLFFYKRSVMSKLSQIGVQHLERVLRL 589  
 435 VLKLLKLEFEDLVSLRCFFVTEQSKSYKTYRKNIWDMKMSIAD-LKETLA 493  
 590 ELSQEVRHHQDTW---LAMPICRLRFIPKPNGLRPIVMSYSGMTRALGRKKAQHTQ 646  
 494 EVOEKEY---EWWKSLGFAPGKRLIPKTTTPRIMFNKI-----VNSDRKTT 541  
 647 RLKTLFSLNVERTKPHLM-----GSLVGMNDIVTWEAFVLVRALDOT 693  
 542 KLTNTYKLN-----SHMLKTLKRMFKDFGFAVFNVDWMKKYEEFVCKWKQVGQ- 594  
 694 PRMYFKAQVTGAYDAIPQGLVGVVANNIRHSESTYCIQYAVVRDSQGVHKSPRRQ 753  
 595 PKLFFATMDIEKCYDSVNRKSLTKLTKLLSSDFWIMTAQILKRNKNIVIDSKNFRKK 654  
 754 VTLSLDLPYWGQFLKHL-----QSDASALRNSVJLEQSI SMNESSS 797  
 655 -----EMKDTFRQFKALLEGQYPTLSVLENEQNDLNA-KTLTIVEAKQRYFKKDN 708  
 798 LFDFFLHFLRHSVVKIGDRCTOCQIGPQSSLSLTLCSLCPGDMENKLFPAEVORD--- 853  
 709 LLOFVINICQYINFNFGFKYKTKGIPQGLCVSSILSSFYVATEESSLGLRDESNMP 768  
 854 -----GILLRFVDVDFLLVPHLDOAKTFTSLVHGVPEYGCMINQKTVNFPVPEGLG 908  
 769 ENPNVNLMLTDDYLLITQENNAVLFEKLIINVSRENGFKFNKKLQTFPLSPSKF- 827  
 909 GAAPYQLPA-----HCLFPWCGLLDDTOTLE-----VFCDSYGAQTSIK 948  
 828 --AKYGVDSVEEQNIYQDYC--DWIGISIDMTKLTALMPNINLRIEGILCTNLNQTK-- 881  
 949 TSLTFQSVFAGKTMKNKLLSVLRK-CHGLFLDLQVNSLQTVCTINIKFILQAYRPHA 1007  
 882 -----KASMLKKLKSFLMNNITHYFRKTTITTEDFANKTLN--KLFTSGCYKTMQ 930  
 1008 CVIQLPDQVRKNLTPFLGLISSQASCCVAILK 1041  
 931 CAKE--YKDFKKNLAMSMDLEVSILYSVTR 962

## RESULT 6

TERT TERTH STANDARD; PRT; 1117 AA.  
 ID TERT TERTH  
 AC 077448;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Telomerase reverse transcriptase (SC 2.7.7.-) (Telomerase catalytic subunit) (Telomerase subunit P133).  
 GN TERT.  
 OS Tetrahymena thermophila.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;  
 CC Tetrahymena; Tetrahymena.  
 CW NCBI\_TaxID=5911;  
 RN [1]

## RP SEQUENCE FROM N.A.

RC STRAIN-B7;  
 RX MEDLINE=98337940; PubMed=9671703;  
 RA Bryan T.M., Sperger J.M., Chapman K.B., Cech T.R.;  
 RT "Telomerase reverse transcriptase genes identified in Tetrahymena  
 thermophila and Oxytricha trifallax";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8479-8484 (1998).  
 RN [2]

## RP SEQUENCE FROM N.A.

RX MEDLINE=98337941; PubMed=9671704;  
 RA Collins K., Gandhi L.;  
 RT "The reverse transcriptase component of the Tetrahymena telomerase  
 ribonucleoprotein complex";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8485-8490 (1998).  
 CC -!- FUNCTION: TELOMERASE IS A RIBONUCLEOPROTEIN ENZYME ESSENTIAL FOR  
 THE REPLICATION OF CHROMOSOME TERMINI IN MOST EUKARYOTES. IT  
 ELONGATES TELOMERES. IT IS A REVERSE TRANSCRIPTASE THAT ADDS  
 SIMPLE SEQUENCE REPEATS TO CHROMOSOME ENDS BY COPYING A TEMPLATE  
 SEQUENCE WITHIN THE RNA COMPONENT OF THE ENZYME.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- SIMILARITY: BELONGS TO THE REVERSE TRANSCRIPTASE FAMILY.  
 CC TELOMERASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF062652; AAC39135.1; -.  
 DR EMBL; AF061284; AAC39140.1; -.  
 DR PIR; T14891; T14891.  
 DR InterPro; IPR000477; RVTse.  
 DR InterPro; IPR003545; Telomerase\_RT.  
 DR Pfam; PF00078; rvt; 2.

DR PRINTS; PR01365; TELOMERASERT.  
 KW Transferase; RNA-directed DNA polymerase; Telomere; Nuclear protein;  
 KW DNA-binding.  
 SQ SEQUENCE 1117 AA; 133317 MW; E5AF15E86B0F0CD8 CRC64;

Query Match 6.8%; Score 401.5; DB 1; Length 1117;

Best Local Similarity 21.2%; Pred. No. 1.7e-21;  
 Matches 173; Conservative 147; Mismatches 358; Indels 139; Gaps 28;

QY 296 SDSLSSGVCCCHKPSSSTLLSPRONAFQLRPFIEHRFLYSRGDQGE----- 344  
 DB 190 NFNFMNGKATSNNNQNNANLSNKKQENQIYPIQSQIFYNHMGREPGVPKSPFFNY 249  
 QY 345 -RLNPSFLSNLQPNLTGARLVEIIFLGSPPRT--SGPLCRTHLSRRY--WQMRPLF- 398  
 DB 250 SEIKKGFGFKVQIEKLGQRQINSKDKIPQHPQTIKKTLKKEYQ-SKNFSCQBERDLFL 308  
 QY 399 ---QQLLVNHAECQYVRLLRSHCRF-----RTANQQVTDALNTSPPH 437  
 DB 309 EFTKIVQNFNINFNLLKFKCKLPENYQSLKSQVKQIVQSEKNANQQSCENLFS--- 355  
 QY 438 LMDLLEHSSPWQYVGFRLACLCKVWSASLWGTGRHNRFRFFKLFISLGYKGLSLQE 497  
 DB 366 ---LYDTEISYKQITNFLRQIQNCVNPQLLG-KGNFKVLEKLYEFVQMKRFENQKVL 421  
 QY 498 LMWKKVEDCHWLSSPGKDRVPAAEH-RLBERLATFLWMDTVVOLLRSFFYITES 556  
 DB 422 YICFMDVDFVEMFVDLKNQKFTQKRKIYSDKKIKLGLIIVIPVIRNFYITEK 481  
 QY 557 TPQKRLFFYKRSVMSKLSQIGVQHLERVLRLSLOEVRHH---QDTWLAMPICRLRF 613  
 DB 482 HKEGQIYFVRKPIWKLVSCLTI-----VKLEENLEKVEEKLIPDSFOKYPOGKLRI 535  
 QY 614 IPKPNGLRPIVMSYSGMTRALGRKKAQHTQSLKTLF--SMLNERTKH--PHLMGSS 669  
 DB 536 IPKGSFRPIMTFL-----BKDKQKNIKLNQILMDSQLVFNRLKMLGQKIGYS 586





Y 679 TWRAFLRV--RALDQTRPMYFVKADVTGAYDAIPQGLVVEVWAMIRHSESTYCIQY- 735  
 b 505 R-KEFKQRLKKFNVLPELYFMKFDVKSCYDSIPRMECNRIKLDALKNENGFFVRSQYF 564  
 Y 736 -----AVVRDSQGVHKSFRQVTLSDLPYMGQFLKHLQDSASALNSVV 784  
 b 565 FNTNTGVKLNVNVRASVPKPYELIDNVTV-----HLSNQDV----- 504  
 Y 785 IEQSIISMNESSSLDFFLHFLRHSVVKIGDRCYTCQGGIPQSSSLTLLCSLCFGDM-- 842  
 b 605 -----INVWEIPI-----KTALWVEDKCYIREDEGLFGSSLSAPIVDLVYDILLE 650  
 Y 843 -ENKLPFAEVRDGLLRVDFLLVTHLDOAKTFLSTLVHGVPEYCCMINLOXTVVPF 901  
 b 651 FYSEFKAFSPQDILLKLADFLIISTDQOQVINKKLMGFGQKYNKANRDKIL----- 706  
 Y 902 VEPGTGGAAPYQLPAHCLFPCWGILLDTQLEVFCDYSGVATOSIKTSLTFOSVFKAG 961  
 b 707 -----AVSQSDDDTVIQFCAMGHIFVKELEVKHSS-----TWNHFIRSKSK 750  
 Y 962 TWNRKLSVRLKCHGLFLDQVNSLQVNCINI-----YK-----IFLQAYR 1004  
 b 751 GFERSIALFNTIRISYKIDTNLSTNTVLMQIDHVVKNISECYKSAFKDLSINTQNMQ 810  
 Y 1005 RHACVIQLPFQDORV 1018  
 b 811 FHS-----FLORI 818

RESULT 8  
 ASL HUMAN  
 D TAST HUMAN STANDARD; PRT; 778 AA.  
 C Q12615;  
 T 30-MAY-2000 (Rel. 39, Created)  
 T 30-MAY-2000 (Rel. 39, Last sequence update)  
 T 28-FEB-2003 (Rel. 41, Last annotation update)  
 E Trophinin-associated protein (Tastin) (Trophinin-assisting protein).  
 N TROAP.  
 S Homo sapiens (Human).  
 C Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 X NCBI\_TaxID=9606;  
 X [1]  
 P SEQUENCE FROM N.A.  
 X MEDLINE=95278733; PubMed=7758945;  
 A Fukuda M.N., Sato T., Nakayama J., Klier G., Mikami M., Aoki D.,  
 A Nozawa S.;  
 T "Trophinin and tastin, a novel cell adhesion molecule complex with  
 potential involvement in embryo implantation.";  
 L Genes Dev. 9:1199-1210(1995).  
 C -|- FUNCTION: COULD BE INVOLVED WITH BYSTIN AND TROPHININ IN A CELL  
 C ADHESION MOLECULE COMPLEX THAT MEDIATES AN INITIAL ATTACHMENT OF  
 C THE BLASTOCYST TO UTERINE EPITHELIAL CELLS AT THE TIME OF THE  
 C EMBRYO IMPLANTATION.  
 C -|- SUBUNIT: DIRECTLY BINDS BYSTIN, AND INDIRECTLY TROPHININ.  
 C -|- SUBCELLULAR LOCATION: Cytoplasmic.  
 C -|- TISSUE SPECIFICITY: STRONG EXPRESSION AT IMPLANTATION SITES. WAS  
 C EXCLUSIVELY LOCALIZED TO THE APICAL SIDE OF THE  
 C SYNCYTOTROPHOBLAST. ALSO FOUND IN MACROPHAGES.  
 C -----  
 C THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 C between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 C the European Bioinformatics Institute. There are no restrictions on its  
 C use by non-profit institutions as long as its content is in no way  
 C modified and this statement is not removed. Usage by and for commercial  
 C entities requires a license agreement (See <http://www.isb.ch/announce/>  
 C or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 C -----  
 C ENBL; U04810; AAA79333.1; --  
 R FIR; I38487; I38487.  
 R Genew; HGNC:12327; TROAP.  
 R MIM; 603872; --

DR GO; GO:0005737; C:cytoplasm; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 KW Cell adhesion; Repeat.  
 FT DOMAIN 504 687 CYS-RICH.  
 FT DOMAIN 516 647 4 X 33 AA APPROXIMATE TANDEM REPEATS.  
 FT REPEAT 516 548 1.  
 FT REPEAT 549 581 2.  
 FT REPEAT 582 614 3.  
 FT REPEAT 615 647 4.  
 SQ SEQUENCE 778 AA; 83758 MW; 52BAB17165672AC0 CRC64;  
 Query Match 22.2%; Score 131; DB 1; Length 778;  
 Best Local Similarity 22.4%; Pred. No. 0.1;  
 Matches 107; Conservative 44; Mismatches 167; Indels 160; Gaps 26;  
 Y 2 TRAPRCVAVSLRSRYREVWPLATFVRRLGPEGRRLLVOPGPKYRILVA-----QC 54  
 Db 358 TPESTPRVQ-----AQLRGVSPQS-----CSEDPALFWEQVAVRLFDQESC 401  
 Y 55 LVCMHWSQPP---PADLSFHOVSSLKELVAR---VVQRLCERNERNVLAFGFELLNEAR 108  
 Db 402 IRSLEGSKPPVATPSGPHSNRTPSLOEVKIQRIGILOQLLQEQVEGLVGQCVPLN--- 458  
 Y 109 GGP--PMAFTSSVRSYLPNTVETLRVSGAMLLLSRVGDDLLVYLLAHCALYLLVPPSC 166  
 Db 459 GGSLLDMVELQLLLEISRTLNATEHNSGT-----SHLPGLLKHSGL-----PKPC 504  
 Y 167 AYQVCGSPLYCICATTDIW-----SYSASVPRTPVGRNFTN-L 205  
 Db 505 LPEEGEP--QCPAPBPBPPEAFCSRSEPEIPEPSLOEQLEVPPEYPPAEP--RLESCC 560  
 Y 206 RFLQIKSSRQE-----APKPL-----ALPSRGTKEHLSLTSTSVPSAKK 246  
 Db 561 RSEPEIPESSRQEQLEVPPEPCPAEPPLSESYCRIPPESSRQEQ-----EYVE--- 612  
 Y 247 ARCPVPRVEGHRQVLPTPSGKSWVPSPARSPVPTAEKDLSSKGKVDLSLSGV--- 304  
 Db 613 ----PCPAPFPGP---LQPTQSCQSGPPGFC-----PRVELGASEPTLEHRSLESSLP 660  
 Y 305 CCKH-KPSSTSL-----SPPRQNAQLRPFETRHFYLSRGDQGERLNPSFLSN 354  
 Db 661 CCSQWAPATSLFSSQHPCLCASPFCISLQSLRP-----PAGQAG-----LSN 703  
 Y 355 LQPNLTGARLV-----EIPLGRPRTSGP--LCRTHRLSRRYQ 393  
 Db 704 LAPRTLALRESLSCLTAIRCFHEARLDDECAFYTSRASPSGPTRYCTNPVATLLEWQ 761

## RESULT 9

DYHC PARTE STANDARD; PRT; 4540 AA.  
 ID DYHC PARTE  
 AC Q27171;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Dynein heavy chain, cytosolic (DYHC).  
 GN DHC-8.  
 OS Paramaecium tetraurelia.  
 OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Peniculida;  
 OC Paramaecium.  
 OX NCBI\_TaxID=5888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Stock 51;  
 RX MEDLINE=96157890; PubMed=8589455;  
 RA Kandi K.A., Forney J.D., Asai D.J.;  
 RT "The dynein genes of Paramaecium tetraurelia: the structure and  
 RL Mol. Biol. Cell 6:1549-1562(1995).  
 CC -|- FUNCTION: Cytoplasmic dynein acts as a motor for the intracellular  
 CC retrograde motility of vesicles and organelles along microtubules.  
 CC Dynein has ATPase activity; the force-producing power stroke is  
 CC thought to occur on release of ADP.

497	QY	ELM--WKKVVEDCHW-----LRSSPGKQVPAERLRERILATP--LF-----WLMDT 541
514	Db	RAKQYIAETIESVEATITTNLRDLSGS-----ASSAKEMFRILAKFNKLFSRPRIKGAIOE 569
542	QY	YVWQLLRSFFVITES-----TFQK--NRLFFYRK-----SVMSKLSQISGVROHLE 584
570	Db	YQSLLATVHKDIQSLQNKFKETQYQKSNRSLASARDIPTLSGFFVINSKQLQIRLQKVMQ 629
585	QY	RV-----RURELSQ-----BEVVRHODTWLAMPICHLRF 613
630	Db	KVEQILGPQWAEEDTGCKCKEMGETFERILDSGPALEDWKQEIHNH-----KA 678
614	QY	IPKENGRLPTVNNYSYMGTRALGRRRQAOCHFTQRLKTLFSLMNVERTKHPHLMGSSVLGM 673
679	Db	VSQNEKLFVVTERRGLEIRVNVYEKLSQLF-KEVRNLSNM-----KTKVPYSISHIA--- 730
674	QY	NDIYRTWRAFLVLRALDQTPRMFYFKVADVTGAYDAIQGKLVEVVANMRHSSTYCIIR 733
731	Db	NDAKASY-PFAL---SLOSLSHTYI---QITSQLNA-KSAKIV----- 765
734	QY	QYAVVRDSDSQGVHKFSFRQVTTLSDLQYMQQFLKHL----- 771
766	Db	AAIRKEVQLQICGGFNLYLTHKTLQGYVKKFTDKVFELEQAVNGLNERIGQIESLCE 823
772	QY	QSDASALRNSVWIEQISWNSSSLDFDFLHFRHSVVKIGDRCYTCOCGIP 825
824	Db	AMKTCPVDSLADLKDIQEVDSLGFNFSN-----LHIWIQIDK----- 864
826	QY	QGSLSLTLLCSLFCFGMKNK-----FAEVQRDGLL-----LRFVDDFLVTPHLDOA 873
865	Db	QIESILDRVTVMQKWLQNFQINQKIQERGLVQTVVHELKLDQIIVDPFPVEYA 921
874	QY	KTFLLSTLVHGPBYGCMINLQKTVVN 899
922	Db	KYFWFQBFHRM--IGQICSLPRLVAN 945

RESULT 10

ID	GLND_VIBVU	STANDARD;	PRT;	873 AA.
AC	Q8DEG3;			
DT	15-SEP-2003 (Rel. 42, Created)			
DT	15-SEP-2003 (Rel. 42, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	[Protein-Pii] uridylyltransferase (EC 2.7.7.59) (Pii uridylyl-transferase) (Uridylyl removing enzyme) (Urase).			
DE	GLND OR VV1857.			
OS	Vibrio vulnificus.			
OS	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;			
OX	Vibrionaceae, Vibrio.			
OX	NCBI_TaxID=672;			
RN	[1]			
SEQUENCE FROM N.A.				
RP	STRAIN=CMCP6;			
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H., Choy H.E.;			
RT	"Complete genome sequence of Vibrio vulnificus CMCP6.";			
RT	Submitted (DSC-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: Modifies, by uridylylation or deuridylylation the Pii			
CC	(glnd), regulatory protein [by similarity].			
CC	-!- CATALYTIC ACTIVITY: UTP + [protein-Pii] = diphosphate + uridylyl-[protein-Pii].			
CC	-!- SIMILARITY: Belongs to the glnd family.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

SEQUENCE FROM N.A.
RC STRAIN=El Tor N16951 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fletschmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
RL CC -!- FUNCTION: Modifies, by uridylylation or deuridylylation the PII
RL CC (GlnB) regulatory protein (by similarity).
RL CC -!- CATALYTIC ACTIVITY: UTP + [protein-PII] = diphosphate + uridylyl-
RL CC [protein-PII].
RL CC -!- SIMILARITY: Belongs to the glnD family.
RL CC -----
RL CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RL CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RL CC the European Bioinformatics Institute. There are no restrictions on its
RL CC use by non-profit institutions as long as its content is in no way
RL CC modified and this statement is not removed. Usage by and for commercial
RL CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
RL CC or send an email to license@isb-sib.ch).
RL CC -----
DR EMBL; AE004298; AAP95406.1; -
DR PIR; E82097; E82097.
DR TIGR; VC2262; -.
DR HAMAP; MF_00277; -.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR006574; HD.
DR InterPro; IPR003607; Met_phosphohydro.
DR InterPro; IPR002934; NTP_transf.
DR Pfam; PF01842; ACT; 2.
DR Pfam; PF01966; HD; 1.
DR Pfam; PF01909; NTP_transf_2; 1.
DR SMART; SM00471; Hdcg; 1.
DR Transference; Nucleotidyltransferase; Complete proteome.
SQ SEQUENCE 861 AA; 102094 MW; 8F740653AC26738 CRC64;
Query Match 2.0%; Score 116; DB 1; Length 881;
Best Local Similarity 19.6%; Pred. No. 1.6;
Matches 107; Conservative 63; Mismatches 161; Indels 214; Gaps 27;

```

QY	427	VTDAINTSPPHLMDDLLRHSSPNQVYGFRLACLCKVGSASLWGTENRRERFFKNLKKFIS	486
DB	45	VTD-LVIGRSEYMDLL-LHRL-WQFEGF-----DSLVEVSL-----	79
QY	487	LKGYCK-----LSLQ-----ELMWNKVYEDCHWL-----SSP	514
DB	80	VGGRGRELHPLSDIDLLVLSQPLSEQVANKISOFLTLMLDKLEIGHAVRTVEQCAEI	133
QY	515	GKDRVPAEHRHLRERILATFLPMLMTYVQLLRFFYITESTFOQNLFFYRKSVMSL	574
DB	140	KQADLTATNLQEARLCC-----CESTHRLKQWVTHSESFWPS-----	178
QY	575	OSIGVRGHLRVRRLRSEBVEHHODTWLAMPICRLRPIPKNGLRIPVNMYSMSMTETRA	634
DB	179	-----EIFYQAKVREQERHARYHDTTNYLEPDIK-----STFGGLRDIHTLSM-VARRH	227
QY	635	LQRRK-----QAQHTQ-----LKTLSMLNY-----ERTKHP	663
DB	228	FGATSLYEMSRFGFLTDAEYRELVECODEFLWRFPALHIELKRYDNRTFAHQVQVAKHL	287
QY	664	HLMGSSVLG-----MNDIYRWRAFVLRVLAALQDTPMTVFVKADVTGAYDAIP-----	711
DB	298	GYPEGNGRIEMWKEFFRTL-----RVAELNKMLLKIFDKAILNNGEEAAEAVIIDDDFQ	343
QY	712	-CGKLV-----EVVANNIHH--SESTY-----CIRQVAVVR-----	740
DB	344	RRGNMIBARKPAUFQARPETILDMFLHMASDSTIESVAPATNRQLRTAARRLNKLFLHTLP	403









GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on on: February 4, 2004, 14:09:14 ; Search time 45 Seconds  
(without alignments)  
6434.112 Million cell updates/sec

itle: US-09-042-460-2

erfect score: 5901

quence: 1 MTRAPRCFAVRLSRGYRE.....TILKAADPALSTDRQTIID 1122

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

arched: 830525 seqs, 258052604 residues

otal number of hits satisfying chosen parameters: 830525

inimum DB seq length: 0

aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

atabase :

- 1: sp\_archea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

result No.	Score	Query Match	Length	ID	Description
1	4751	80.5	1128	11 Q9QXZ4	Q9QXZ4 mesocricetu
2	3217.5	54.5	1069	4 Q8NG46	Q8NG46 homo sapien
3	3002	50.9	575	11 Q9JK99	Q9JK99 rattus norv
4	2729	46.2	514	11 Q9R266	Q9R266 mus musculu
5	2397.5	40.6	1191	13 Q9DE32	Q9DE32 xenopus lae
6	2195	37.2	807	4 Q8N6C3	Q8N6C3 homo sapien
7	2125	36.0	795	4 Q8NG38	Q8NG38 homo sapien
8	1373	23.3	524	4 Q8UBR6	Q8UBR6 homo sapien
9	1086.5	18.4	523	4 Q84807	Q84807 homo sapien
10	780.5	13.2	1123	10 Q8S899	Q8S899 arabidopsis
11	777.5	13.2	1123	10 Q8SPU7	Q8SPU7 arabidopsis
12	729	12.4	1259	10 Q8AU13	Q8AU13 cryza sativ
13	686	11.6	1261	10 Q8LKW0	Q8LKW0 cryza sativ
14	550	9.3	104	11 Q8JLM1	Q8JLM1 mus musculu
15	450	7.6	1032	5 Q8XUB3	Q8XUB3 euphorbia cr
16	443	7.5	823	5 Q8SQQ0	Q8SQQ0 encephalito

Q8i8z6 sterkiella  
Q9grc5 paramecium  
Q8i8z7 sterkiella  
Q9p8t3 candida alb  
Q8muq8 paramecium  
Q9p8t2 candida alb  
Q9uns6 homo sapien  
Q9unr4 homo sapien  
Q9r0b3 mus musculu  
Q962f9 cryptospori  
Q9ncp5 giardia lam  
Q8i8z2 plasmodium  
Q8i8z1 streptococc  
Q8eqa2 oceanobacil  
Q95qp3 caenorhabdi  
Q45321 caenorhabdi  
Q8dn10 streptococc  
Q9ael7 streptococc  
Q9ids9 streptococc  
Q8dz79 streptococc  
Q8n5b2 homo sapien  
Q8emx5 oceanobacil  
Q22002 caenorhabdi  
Q8ep73 oceanobacil  
Q8enl6 oceanobacil  
Q8bi59 mus musculu  
Q8i068 arabidopsis  
Q8bi79 mus musculu  
Q9he09 schizosacch

## ALIGNMENTS

### RESULT 1

Q9QXZ4 PRELIMINARY; PRT; 1128 AA.  
AC Q9QXZ4;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Telomerase catalytic subunit.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21240330; PubMed=11342218;  
RA Guo W., Okamoto M., Lee Y.M., Baluda M.A., Park N.H.;  
RT "Enhanced activity of cloned hamster TERT gene promoter in transformed cells";  
RL Biochim. Biophys. Acta 1517:398-409(2001).  
DR EMBL; AF149012; AAF17334.1; -;  
DR InterPro; IPR000477; RVISE.  
DR InterPro; IPR003545; Telomerase\_RT.  
DR Pfam; PF00078; rvt. 1.  
DR PRINTS; PR01365; TELOMERASERT.  
KW RNA-directed DNA polymerase; Transferase.  
SQ SEQUENCE 1128 AA; 128394 MW; 1D4F81249012174E CRC64;

Query Match 80.5%; Score 4751; DB 11; Length 1128;  
Best Local Similarity 80.1%; Pred. No. 0;  
Matches 904; Conservative 90; Mismatches 127; Indels 8; Gaps 4;

QY 1 MTRAPRCFAVRLSRGYREVWPLATFVRLGEGRLVQPGDPKIYRTLVQAQCLVCMHW 60  
DB 1 MPRAPRCFAVRLSRGYREVWPLATFVRLGEGRLVQPGDPKIYRTLVQAQCLVCMHW 60  
QY 61 GSQPPADLSFHQVSSLKELVARVQVQRLCERNERNVLAFGFELLNARGCPPTTSYR 120  
DB 61 DSQPPADLSFHQVSSLKELVARVQVQRLCERNERNVLAFGFELLNARGCPPTTSYR 120



```

121 SYLNTVETIRVSGAMMLLSRVGDDLLVYLHQAALVLLVPPSCAYQVCGSPYQICA 180
121 SYLNSVETIRVSGAMMLLSRVGDDLLVYLHQAALVLLVPPSCAYQVCGSPYQICA 180
181 TTDIWPVSASVPTPRVGNFTNLFLQIKSSSQEAPKPLPSRGTGKHLSTGTS 240
181 TAETWPSVRIYPTPRVGNFTNLFLQIKSSSQEAPKPLPSRGTGKHLSTGTS 240
241 VPSAKKARCYVPRVBEGRHQVLPPTPSGKSVWPSPARSPFVPT---TAEKOLSSGKGVSD 297
241 VPSKAKARCYVPRVBEGRHQVLPPTPSGKSVWPSPARSPFVPT---TAEKOLSSGKGVSD 300
298 LSLSGSVCCGKHKSSSTLSLPPQNAQFQRLPRTETRFYLYSGDGERLNSFLSNLQ 357
301 LSRSGSVCCGKHKSSSTLSLPPQNAQFQRLPRTETRFYLYSGDGERLNSFLSNLQ 360
358 NLGARELVETIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHABQVYVLLRSH 417
361 SLTGARELVETIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHABQVYVLLRSH 420
418 CRFTANQVTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCVKVSGASLWGTREHNR 476
421 CRFTANQVTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCVKVSGASLWGTREHNR 480
477 FFKNLKXFIISLGKYKLSLQELMWMKVEDCHWLRSSPGKDRVPAEHLRERILATLFL 536
481 FFKNLKXFIISLGKYKLSLQELMWMKVEDCHWLRSSPGKDRVPAEHLRERILATLFL 540
537 WLMDTVYVQLLRFFYITESTFQKRLPLFYKSVWSKQSLQISGVROHLRVLRLSQEVE 596
541 WLMDTVYVQLLRFFYITESTFQKRLPLFYKSVWSKQSLQISGVROHLRVLRLSQEVE 600
597 RHQDWTWLMPIICRLRFIPKNGRLIPVNMSSMGTRALGRKQOQHTFQRLKTLFSLN 656
601 RQCEAWPMPICELRFIPKNGRLIPVNMSSMGTRALGRKQOQHTFQRLKTLFSLN 659
657 YERTKPHLMGSSVLGNDIYRTWRAFLVRLRALDQTPMYFVKADVTGAYDAIPQKLV 716
660 YELTKHTNLLGASVLGNDIYRTWRAFLVRLRALDQTPMYFVKADVTGAYDAIPQKLV 719
717 EVANMIRHSESTYCIROYAVRDSOGVHKSFRQVNTLSLQPYMGQFVHLQDSDA 776
720 EVANMIRHSESTYCIROYAVRDSOGVHKSFRQVNTLSLQPYMGQFVHLQDSDA 779
777 SALNSVYVQCSISNMSSSLDFLFLHLSVWVKIGRCYTCQGIPOGSSSLTLCSS 836
780 SALNSVYVQCSISNMSSSLDFLFLHLSVWVKIGRCYTCQGIPOGSSSLTLCSS 839
837 LCFGDMENKLFABVQDGLLRLRVDFDPLVTHLDOAKTFLSTLVHGVPEYGCMLNLOKT 896
840 LCFGDMENKLFABVQDGLLRLRVDFDPLVTHLDOAKTFLSTLVHGVPEYGCMLNLOKT 899
897 VVNPFPVETLGGAAAPYQALPFAHCLFPCWGLLDTQTLFVFCDSYGAQTSIKTSLTFQSV 956
900 VVNPFPVETLGGAAAPYQALPFAHCLFPCWGLLDTQTLFVFCDSYGAQTSIKTSLTFQSV 959
957 FKAGTWKRLSVLRKCHGLFLDLQVNSLQTVNCINIKIFLQAYRFAHCVQLPDPQ 1016
960 FKAGTWKRLSVLRKCHGLFLDLQVNSLQTVNCINIKIFLQAYRFAHCVQLPDPQ 1019
1017 RVKRLNLTFFLGITSSOASCCYALVKVNPQWTLK---ASGSPFPAAJWLCVQAFLLKLA 1073
1020 HVRKNPFAFFLIISNATSCCYSLKVNAGWTLKAGAGSFPFAARWLVCYQAFLLKLA 1079
1074 AHSVYKCLLGLPRLTAQKLCRLKPLBATWTLKAAADPALSTDFQTLTD 1122
1080 GHSVYKCLLGLPRLTAQKLCRLKPLBATWTLKAAADPALSTDFQTLTD 1128

```

RESULT 2

JEN46

CD Q8NG46

PRELIMINARY; PRT; 1069 AA.

```

AC Q8NG46;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Telomerase reverse transcriptase.
GN HTERT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hisatomi H., Nagao K., Hirata H., Hiki K., Kanamaru T.;
RT "Exon 11 deleted variant of the human telomerase reverse
RL transcriptase.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB085628; BAC11010.1; -.
DR InterPro; IPR003545; Telomerase_RT.
DR PRINTS; PR01365; TELOMERASERT.
KW RNA-directed DNA polymerase.
SQ SEQUENCE 1069 AA; 120047 MW; BE1E77A653B1C666 CRC64;

Query Match 54.5%; Score 3217.5; DB 4; Length 1069;
Best Local Similarity 58.4%; Fred. No. 3.7e-263;
Matches 673; Conservative 117; Mismatches 248; Indels 115; Gaps 14;

QY 1 MTRAPRCAPVRSLLRSRYREVPLATFVRRLLGPEGRRLLVOPGDPKIYRTLVQAQCLVCMHW 60
DB 1 MPRAVRCAPVRSLLRSRYREVPLATFVRRLLGPEGRRLLVOPGDPKIYRTLVQAQCLVCMHW 60
QY 61 GQOPPADLSHFQVSSSLKELVARVVQRLCERNVNLAFGPELLNEARGPPMAFTSSVR 120
DB 61 GQOPPADLSHFQVSSSLKELVARVVQRLCERNVNLAFGPELLNEARGPPMAFTSSVR 120
QY 61 DARPPFAAPSPRQVSCLEKELVARVQLCERGANVLAFCGFDLGGARGGPPPEAFTTSVR 120
DB 61 DARPPFAAPSPRQVSCLEKELVARVQLCERGANVLAFCGFDLGGARGGPPPEAFTTSVR 120
QY 121 SYLNTVETIRVSGAMMLLSRVGDDLLVYLHQAALVLLVPPSCAYQVCGSPYQICA 180
DB 121 SYLNTVETIRVSGAMMLLSRVGDDLLVYLHQAALVLLVPPSCAYQVCGSPYQICA 180
QY 181 TTDIWPVSASVPTPRVGNFTNLFLQIKSSSQEAPKPLPSRGTGKHLSTGTS 240
DB 181 TTDIWPVSASVPTPRVGNFTNLFLQIKSSSQEAPKPLPSRGTGKHLSTGTS 240
QY 181 ATQARPPPHAS-GPRRLG-----CERANWNSVREAGVPLCLPAGARRRGGASRS 231
DB 181 ATQARPPPHAS-GPRRLG-----CERANWNSVREAGVPLCLPAGARRRGGASRS 231
QY 241 VPSAKKARCYVPRVBEGRHQVLPPTPSGKSVWPSPARSPFVPTAEKOLSSK 292
DB 241 VPSAKKARCYVPRVBEGRHQVLPPTPSGKSVWPSPARSPFVPTAEKOLSSK 292
QY 232 LPLPKRRRGAAPERTPVCGGSHWHPGTRGSDRGFCVVSPEAR-----PABEATSL 286
DB 232 LPLPKRRRGAAPERTPVCGGSHWHPGTRGSDRGFCVVSPEAR-----PABEATSL 286
QY 293 GKVSLLSLS-GSVCCCKXSPSTLSLSPRQNAFQRLP-FIETRHFLYSRGDGOERLNPSF 350
DB 293 GKVSLLSLS-GSVCCCKXSPSTLSLSPRQNAFQRLP-FIETRHFLYSRGDGOERLNPSF 350
QY 287 GALSGRHSHPSVGRHQHAGPPSTSRPRPMDTPCPVYAEKHFYSSGD-KEQLRPSF 345
DB 287 GALSGRHSHPSVGRHQHAGPPSTSRPRPMDTPCPVYAEKHFYSSGD-KEQLRPSF 345
QY 351 LLSNLQPLNLTGARRLVEIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHAEQY 410
DB 351 LLSNLQPLNLTGARRLVEIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHAEQY 410
QY 346 LLSLRPSLTGARRLVETIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHAEQY 405
DB 346 LLSLRPSLTGARRLVETIIFLGSPPRTSGPLCLTHRLSRVWQMRPLFOQLLVNHAEQY 405
QY 411 VRLLSHCRFRATANQVTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCVKVSGASLWGTREHNR 452
DB 411 VRLLSHCRFRATANQVTDALN-TSPHMLDLRLHSSPMQVYGFRLACLCVKVSGASLWGTREHNR 452
QY 406 GVLLKTHCPLRAA---VTPAAGVCAKPKQSVAAPEEDTDPRRLVQLLRQSSPMQVY 462
DB 406 GVLLKTHCPLRAA---VTPAAGVCAKPKQSVAAPEEDTDPRRLVQLLRQSSPMQVY 462
QY 453 GFRLACLCKVVSASLWGRTHNERFFKNLKFISLGKYKLSLQELMWMKVEDCHWLR 512
DB 453 GFRLACLCKVVSASLWGRTHNERFFKNLKFISLGKYKLSLQELMWMKVEDCHWLR 512
QY 463 GFVRACLARLVPPGLWSGRHNERFLNTRKTFISLGKHAQLSLQELTWKNSVRDCAWLR 522
DB 463 GFVRACLARLVPPGLWSGRHNERFLNTRKTFISLGKHAQLSLQELTWKNSVRDCAWLR 522
QY 513 SPKDRVPAAEHRLRERILATFLEWMDTVYVQLLRFFYITESTFQKRLPLFYKSVWS 572
DB 513 SPKDRVPAAEHRLRERILATFLEWMDTVYVQLLRFFYITESTFQKRLPLFYKSVWS 572
QY 523 SPFGVGVPAAEHRLRERILATFLEWMDTVYVQLLRFFYITESTFQKRLPLFYKSVWS 582
DB 523 SPFGVGVPAAEHRLRERILATFLEWMDTVYVQLLRFFYITESTFQKRLPLFYKSVWS 582
QY 573 KLSQISGVKHLRVLRLSQEVEYRHHQDTWLMPIICRLRFIPKNGRLIPVNMSSMG 632
DB 573 KLSQISGVKHLRVLRLSQEVEYRHHQDTWLMPIICRLRFIPKNGRLIPVNMSSMG 632
QY 583 KLSQISGVKHLRVLRLSQEVEYRHHQDTWLMPIICRLRFIPKNGRLIPVNMSSMG 642
DB 583 KLSQISGVKHLRVLRLSQEVEYRHHQDTWLMPIICRLRFIPKNGRLIPVNMSSMG 642
QY 633 RALGRKQAQHTQRLKTLPSMLNRYETKPHLMGSSVLGNDIYRTWRAFLVRLRALDQ 692
DB 633 RALGRKQAQHTQRLKTLPSMLNRYETKPHLMGSSVLGNDIYRTWRAFLVRLRALDQ 692
QY 643 RTFRKRAERLTSRVKALFSLVNYEARPPGLLGASVLGLDDIHRAWRTFVLRVRAQDP 702
DB 643 RTFRKRAERLTSRVKALFSLVNYEARPPGLLGASVLGLDDIHRAWRTFVLRVRAQDP 702

```

[illegible]

```

1 RESULT 3
2 JJK99
3 ) PRELIMINARY; PRT; 575 AA.
4
5 09JK99;
6 01-OCT-2000 (TREMBLrel. 15, Created)
7 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
8 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
9 Telomerase catalytic subunit (Fragment).
10 Rattus norvegicus (Rat).
11 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
12 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
13 NCBI_TaxID=10116;
14 {1}
15 SEQUENCE FROM N.A.
16 Wong S., Gao S., Xu X., Yu H.;
17 "Rat telomerase catalytic subunit, rTERT.";
18 Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
19 EMBL; AF247818; AAF62177.1; -.
20 InterPro: IPR000477; RVTse.
21 InterPro: IPR003545; Telomerase_RT.
22 Pfam: PF00078; rvt; 1.
23 PRINTS; PR01365; TELOMERASERT.
24 RNA-directed DNA polymerase; Transferase.
25 1
26 NON TER
27 SEQUENCE 575 AA; 65672 MW; F80C8-BD7F6A91A3 CRC64;

```

Db	121	SSVLGMNDIVRTWFAFLVRALDQTPRMYFKYADVTGAYDAIPQGLVEVVANNIRHSE	180
Qy	728	STYCIROYAVVRSDSQCVHKSPRRQVTTLSLDLPYMGQFLXHLQSDASALRNSVVIQ	787
Db	181	STYCIROYAVVRSDSQCVHKSPRRQVTTLSLDLPYMGQFLXHLQSDASALRNSVVIQ	240
Qy	789	SI-SMNESSSSLFDFLHFLRHSVVKIGDRCYTCQOGIPQGSLSLTLLCSLFCGMENKLF	847
Db	241	SI-SMNESSSSLFDFLHFLRHSVVKIGDRCYTCQOGIPQGSLSLTLLCSLFCGMENKLF	300
Qy	848	AEVQDQGLLLRFVDDFLVTPHLDDQAKTFLSTLVHGVPYEGCMINLQKTVNFPVEPGLT	907
Db	301	AEVQDQGLLLRFVDDFLVTPHLDDQAKTFLSTLVHGVPYEGCMINLQKTVNFPVEPGLT	360
Qy	908	GGAAFPYQLPAHCLFPWCGLLDQTTLVEFCDYSGYAQTSIKTSLTFQSFVFKAGTMRNKL	967
Db	361	GGAAFPYQLPAHCLFPWCGLLDQTTLVEFCDYSGYAQTSIKTSLTFQSFVFKAGTMRNKL	420
Qy	968	LSVLRLKCHGFLFDLQNSLQTVGCINIKYIFLLOAYRHACVQIQLPFOQRVRKMLTFPLG	1027
Db	421	LSVLRLKCHGFLFDLQNSLQTVGCINIKYIFLLOAYRHACVQIQLPFOQRVRKMLTFPLG	480
Qy	1028	IISSQASCCYAILKVKQPGMTLTKASGSFPPEAAHMLCYQAFLLKLAASHSVYKCLLGPRL	1087
Db	481	IISSQASCCYAILKVKQPGMTLTKASGSFPPEAAHMLCYQAFLLKLAASHSVYKCLLGPRL	540
Qy	1088	TAQKLLCRKLEPATWTLKAAADPALSTDFQTILD	1122
Db	541	TAQKLLCRKLEPATWTLKAAADPALSTDFQTILD	575

```

RESULT 4
Q9R266      PRELIMINARY;      PRT;      514 AA.
ID          AC
Q9R266;
DT          01-MAY-2000 (TRENBERel. 13, Created)
DT          01-MAY-2000 (TRENBERel. 13, Last sequence update)
DT          01-MAY-2000 (TRENBERel. 13, Last annotation update)
DE          Telomerase reverse transcriptase (Fragment).
          TERT.
GN          GN
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC          NCBI_Taxid=10090;
RX          [1]
RP          SEQUENCE FROM N.A.
RX          GREENBERG-99144726; PubMed=10022128;
RX          Lichtsteiner R.A., O'Hagan R.C., Deng H., Xiao Q., Hann S.R., Adams R.R.,
RA          Lichtsteiner S., Chin L., Morin G.B., Depinho R.A.:
RT          "Telomerase reverse transcriptase gene is a direct target of c-Myc but
RT          is not functionally equivalent in cellular transformation.";
RL          Oncogene 18:1219-1226(1999).
DR          ENBL; AF121949; AAD24465.1; -.
KW          RNA-directed DNA polymerase.
FT          NON TER      514
SQ          SEQUENCE      514 AA; 58382 MW; 3096599776D9BBFD CRC64;
          Query Match      46.2%; Score 2729; DB 11; Length 514;
          Best Local Similarity 100.0%; Pred. No. 2.9e-222;
          Matches 514; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

01-OCT-2002 (TREMBlrel. 22, Created)
01-OCT-2002 (TREMBlrel. 22, Last sequence update)
01-MAR-2003 (TREMBlrel. 23, Last annotation update)
Beta and gamma deletion isoform of telomerase reverse
transcriptase.
HTEPT.
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
TISSUE=Stomach cancer;
Hisatomi H., Nagao K., Kanamaru T., Sumida H., Hirata H., Yamamoto M.,
Kazumasa H.;
"Both beta and gamma deletion isoform of human telomerase reverse
transcriptase.";
Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
EMBL; AB086379; BAC11014.1; -.
InterPro; IPR003545; Telomerase_RT.
PRINTS; PR01365; TELOMERASERT.
SEQUENCE 807 AA; 90225 MW; 199664460CE6D763 CRC64;

Query Match 37.2%; Score 2195; DB 4; Length 807;
Best Local Similarity 58.3%; Pred. No. 9.7e-177;
Matches 462; Conservative 88; Mismatches 191; Indels 52; Gaps 12;

1 MTRAPCPAVRSLLRSRVREVMPLATFVRLGPEGRRVLQPGDKIVRTLVAOCLVCMHW 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MPRAPCPAVRSLLRSRVREVMPLATFVRLGPEGRRVLQPGDKIVRTLVAOCLVCMHW 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

61 GSOPPPADLSFHQVSSLSKELVARVQRLCERNVLAFFGELLNEARGGPPMAFTSSVR 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 DAREPPAAFSFQVSLKELVARVQRLCERNVLAFFGELLNEARGGPPMAFTSSVR 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

121 SYLPTNTVITLRSVGNMLLSRVGGDLLVLLAHCALYLLVPSCAYQVCGSPLYQICA 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 SYLPTNTVITLRSVGNMLLSRVGGDLLVLLAHCALYLLVPSCAYQVCGSPLYQICA 180
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

181 TTDIWPSVASYRPTRPVGRNFTNLRFLOQIKSSROEAPKPLALPSRGTKRHLSTSTS 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 ATQARPPPHAS-GPRRLG-----CERAWNSVREAGVPLGLPAGARRRGGSASRS 231
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

241 VPSAKKARCYVPRVEEGP-----HRQVLTPSGKSW-VPSPARPEVPTAEKLSK 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
232 LPLPKRPRGAAPERTPVQGGWAHPGRTGSDRGFCVVSAPAR-----PABEATSLE 286
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

293 GKVSDDLSS-GSVCKHKPSSTLSPPRQNAFQLRP-FIETRHFLYSRGDOERLNPSP 350
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
287 GALSGTRHSHPSVGRQHHAGPSTSRPRPNDTPCPPVYAEKHFLYSSGD-KEQLRPSF 345
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

351 LLSNLOPNTGARRLVEIIFLGSRRPTSGPLCRTHRLSRRYQWQMRPLFQOLLVNHACQY 410
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
346 LLSSLRPSLTGARRLVEIIFLGSRRPTSGPLCRTHRLSRRYQWQMRPLFQOLLVNHACQY 410
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

411 VLLRSHCHCFRTANQOVTDAL-----NTSPHMLDLLLRLHSSPWQVY 452
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
406 GVLLKTHCHPLRAA---VTPAAGVCAREKPGQSVAAPEEDTDPRLLVQLLRQHSPPQVY 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

453 GFVRACLCKVWSASLWGTNRHRRFPKLLKFIISLGKYGKLSLOELMMKMWEDCHWLRS 512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
463 GFVRACLRLVPPGLWGRNRRERFLNFKFIISLGKHAKLSLQELTWGMSVRCAMLR 522
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

513 SPQKDRVPAEHLRLREIATFLWMDTVYVQLLRSPFYITESTFOKRLFFYFKVWS 572
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
523 SPQGVGVPAEHLRLREIATFLWMDTVYVQLLRSPFYITESTFOKRLFFYFKVWS 582
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

573 KLOSIGVROHLEVRLEISQEBEHHQDTWAMPICRLRIPKENGRLPIVNNYSNGT 632
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
583 KLOSIGVROHLEVRLEISQEBEHHQDTWAMPICRLRIPKENGRLPIVNNYSNGT 642
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

633 RALGRKQAOHFTQRLKTLFSLMNTERTKHPHLMGSSVLGMNDIYRTWRAFLVRALDQ 692
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

b 463 GFVACLRRLVPPGLWGRHNRERFLNRTKFKFISLGHAKLSQELTWKNSVRDCAWLRR 522
y 513 SPGKDRVPAARHLRERLATLFLWMDTYVQLRFFYITESTFQKRLFFYRKSVM 572
d 523 SPGVCVPAARHLRERLATLFLWMDTYVQLRFFYITESTFQKRLFFYRKSVM 582
y 573 KLOSICVQHLEVRVRLSQRVHHQDTWAMPICRLRFIPKPNGLRPIVNMYSMT 632
d 583 KLOSIGIRQHLEVRVRLSQRVHHQDTWAMPICRLRFIPKPNGLRPIVNMYSMT 642
y 633 RALGRKQAQOFTORLTKLTFMLNRYERTKPHLMGSSVLQGNNDIYRTWRAFLVRALDQ 692
d 643 RTFREKEARLTSRVKALFVNLVYERARRPGLGASVLGLDDIHRWRTFVLVRAQDP 702
y 693 TPRMTFVADVTGAYDAIPQOKLVEVNVANMRHSESTYCIQVAVVRDRSQGVHKSFR 752
b 703 PPLEYFVK-----DRLTEVIAIHK-PONTYCVRRYAVVQKAHGHVRKAFKS 749
y 753 QVTTLSDLQPYMG 765
b 750 HV-----LRPVP 757

RESULT 8
9UBR6 PRELIMINARY; PRT; 524 AA.
D QUBR6
T QUBR6;
T 01-MAY-2000 (Tremblrel. 13, Created)
T 01-MAY-2000 (Tremblrel. 13, Last sequence update)
T 01-MAY-2000 (Tremblrel. 13, Last annotation update)
E Telomerase reverse transcriptase (Fragment).
TERT.
S Homo sapiens (Human).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
X NCBI_TaxID=9606;
N [1]
P SEQUENCE FROM N.A.
MEDLINE=99140777; PubMed=10022128;
X Greenberg R.A., O'Hagan R.C., Deng H., Xiao Q., Hann S.R., Adams R.R.,
A Lichtsteiner S., Chin L., Morin G.B., DePinho R.A.;
T "Telomerase reverse transcriptase gene is a direct target of c-Myc but
T is not functionally equivalent in cellular transformation."
L Oncogene 18:1219-1226 (1999).
N [2]
P SEQUENCE FROM N.A.
MEDLINE=99140777; PubMed=9988278;
X Wu K.J., Grandori C., Anacker M., Simon-Vermot N., Polack A.,
A Lingner J., Dalla-Favera R.;
T "Direct activation of TERT transcription by c-MYC."
T Nat. Genet. 21:220-224 (1999).
L R EMBL; AF121948; RAD24464.1; -.
W EMBL; AF114847; RAD17210.1; -.
R RNA-directed DNA polymerase.
T NON_TER 524
Q SEQUENCE 524 AA; 57932 MW; 5F47DEFD01832B1B CRC64;

Query Match 23.3%; Score 1373; DB 4; Length 524;
Best Local Similarity 56.1%; Pred. No. 2.1e-107;
Matches 304; Conservative 50; Mismatches 142; Indels 46; Gaps 10;

y 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPGRLVOPGDPKIVRTLVAOCLVCMHW 60
b 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPGRLVOPGDPKIVRTLVAOCLVCMHW 60
y 61 GSQPPADLSFHQVSSLKELVARVQRLCERNERNVLAFCGELLNARGGPPMAFTSSVR 120
b 61 DARPPAPAFRQVSCVCLKELVARVQRLCERGAKNVLAFCGELLNARGGPPMAFTSSVR 120
y 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLVPSCAYQVCGSPLYQICA 180
b 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLVPSCAYQVCGSPLYQICA 180
y 181 TTDIWPVSASVRYPTREVGNTNLFLOQIKSSRQAPKPLALPSRGTKXHLSTSTS 240
b 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGLPAPGARRRGGSARS 231

```

```

QY 181 TTDIWPVSASVRYPTREVGNTNLFLOQIKSSRQAPKPLALPSRGTKXHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGLPAPGARRRGGSARS 231
QY 241 VPSAKKARCVPVPRVEEGP-----HROVLTPSGKSW-VSPSPARSPEVPTAEXDLSSK 292
DB 232 LPLPKPRRGAAPERTPTVGQGSMAHPCGTRGSPDRGFCVVSPAR-----PABEATSLE 286
QY 293 GKVSDDLSSL-GSVCCCHKFSSTLSLSPRONAFQLRP-FIETRHFLYSGDQOERLNSF 350
DB 287 GALSGRTHSPSVGRCHAGPPSTSRPPRPWDTPCPPVYAEATKFLYSGD-KEQLRPSF 345
QY 351 LLNLQPNLTGARRLVEIIFLGSRRPTSGPLCRTRLRSRYWQMRPLFQOLIVNHAECQY 410
DB 346 LLSLRPLSLTGARRLVETIFLGSRRPWPOTPRRLPRLPQRYWQMRPLFLELLGNHAQCPY 405
QY 411 VLLRSRCRFTANQVTDAL-----NTSPHMLDLLRLHSSFPQVY 452
DB 406 GVLLKTHCPLEAA---VTPAAGVCAREKPGQSVAAPEEBDTPRLVLQRLRHSSFPQVY 462
QY 453 GFLRACLCKVVSASLWGTNRNRRFPKLLKFKISLGVKLSLQELMWKVKVEDCHWLS 512
DB 463 GFVACLRRLVPPGLWGRHNRERFLNRTKFKFISLGHAKLSQELTWKNSVRDCAWLRR 522
QY 513 SP 514
DB 523 SP 524

RESULT 9
O94807 PRELIMINARY; PRT; 523 AA.
ID O94807
AC O94807;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE Telomerase transcriptase (Fragment).
GN HTER.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takakura M., Kyo S., Kanaya T., Hirano H., Takeda J., Yutsudo M.,
RA Inoue M.;
RT "Cloning and characterization of human telomerase catalytic subunit
RT (hTERT) gene promoter."
RL Cancer Res. 0:0-0 (1999).
DR EMBL; AB016767; BAA74724.1; -.
FT NON_TER 523
SQ SEQUENCE 523 AA; 56555 MW; 8FDE562DDECC93DA CRC64;

Query Match 18.4%; Score 1086.5; DB 4; Length 523;
Best Local Similarity 48.7%; Pred. No. 3.9e-83;
Matches 269; Conservative 45; Mismatches 171; Indels 67; Gaps 16;

QY 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPGRLVOPGDPKIVRTLVAOCLVCMHW 60
DB 1 MTRAPRCVAVRLLSRVREYVWPLATFVRRLGPGRLVOPGDPKIVRTLVAOCLVCMHW 60
QY 61 GSQPPADLSFHQVSSLKELVARVQRLCERNERNVLAFCGELLNARGGPPMAFTSSVR 120
DB 61 DARPPAPAFRQVSCVCLKELVARVQRLCERGAKNVLAFCGELLNARGGPPMAFTSSVR 120
QY 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLVPSCAYQVCGSPLYQICA 180
DB 121 SYLNTVETLRVSGAWMLLSRVGDDLLVLLAHCALVLLVPSCAYQVCGSPLYQICA 180
QY 181 TTDIWPVSASVRYPTREVGNTNLFLOQIKSSRQAPKPLALPSRGTKXHLSTSTS 240
DB 181 ATQARPPPHAS-GPRRLG-----CERANWHSVREAGVPLGLPAPGARRRGGSARS 231

```

241 VPSAKARCYVPVREBGP-----HQVLPDSGKSW-VPSAPRSPVPTAEKOLSK 292  
 232 LFLPKPRRGAAPERTEVQGGWAHFGTRTGRSGDRGFCVVSAPR-----PAEEATISLE 286  
 293 GKVSLSLS-GSVCCCKHKPSSSTLSLPPRQNAFOLRP-FIETRHFLYSRGDGOERLNPSF 350  
 287 GALSOTRSHSPVSGQHAGPSSISRPWPDTCPPPYAETHFLYSSGD-KEQLRPSP 345  
 351 LLSNLOPNLTGARRLVEIFLGSPPRTSG-----PLC--RTHLRSSRYWQMRPLFOQLLV 403  
 346 LLSLRPLTLAGRLVETIFLGSPPCGQLPAGCPACPSATGKCGPCFWSCLGTTTSAPT 405  
 404 NHAECQVRLRSHCRFTANQOVTDALNTSPHMLDLLLHSS---PWQVYGFRLACLC 460  
 406 G---CSSRRTAR--CELASPPQVPS-VPGRSFRLWPPRRRTGTPVAM-----CSC 451  
 461 KVSASLWGT-----HNERFFKNLKKFISLKGKLSLOELMWKM 502  
 452 SASAPLAGVRLRAGLPAPAGAPRPLGLQAORRRFLRNTKFFISLGHAKLSLOELTKM 511  
 503 KVEDCHWLRRSP 514  
 512 SVRDCAWLRRSP 523

SULT 10  
 SE99 PRELIMINARY; PRT; 1123 AA.  
 Q9SE99;  
 Q9SE99;  
 01-MAY-2000 (TREMELrel. 13, Created)  
 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
 Telomerase reverse transcriptase catalytic subunit.  
 TERT.  
 Arabidopsis thaliana (Mouse-ear cress).  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 NCBI\_TaxID=3702;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=Landberg erecta;  
 MEDLINE=20079564; PubMed=10611295;  
 Fitzgerald M.S., Rika K., Gao F., Ren S., McKnight T.D., Shippen D.E.;  
 "Disruption of the telomerase catalytic subunit gene from Arabidopsis  
 inactivates telomerase and leads to a slow loss of telomeric DNA.";  
 Proc. Natl. Acad. Sci. U.S.A. 96:14813-14818(1999).  
 EMBL; AF172097; AAD54276.1; --  
 InterPro; IPR000477; RVTse.  
 InterPro; IPR000215; Serpin.  
 InterPro; IPR003545; Telomerase\_RT.  
 Pfam; PF00078; rvt; 1.  
 PRINTS; PRO1365; TELOMERASERT.  
 PROSITE; PS00284; SERPIN; 1.  
 RNA-directed DNA polymerase; Transferase.  
 SEQUENCE 1123 AA; 130551 MW; 44248295817B7F6 CRC64;

Query Match 13.2%; Score 780.5; DB 10; Length 1123;  
 best Local Similarity 24.1%; Pred. No. 1e-56;  
 matches 290; Conservative 208; Mismatches 435; Indels 269; Gaps 42;  
 38 LVQPGDPKIXRTLVQAQVCMHWSGPPADLSFHOVS--SLKELVARVVQRL--CERN 92  
 60 LRSDDPHYRKLHRCFVUHL--EQTPPL-LDFSPTSWWSQREIVERIEMQSGCDC- 115  
 93 ERNVAFOGELLNEARGPPMAFTSVRSYLPNTVETLRVSGAMLLLSKRGDDLLVYL 152  
 116 -QNVICARYD-----KYDQSSPILELTSSSWEFLLKRVGHDVAVYL 156  
 153 LAHCALYLLVPPSCAYQVCGSFLYQICNTTDWPSVSASVPRTPVGRNTNLRFLQOIK 212  
 157 LQOTSIFLPLGLKHKHQVSGGPL-----CIKHKRTLSVHEN----- 192

213 SSSRQERAPKPLALPSRGTQKH-LSLTSTSVPSAKKARCYVPVREBGPVPTAEKOLSK 271  
 193 KRKEDDNNVQP-----PTKQWLSSAVDDCPKDDSATITPIVGEDVDQHKRKKTKRSRI 246  
 272 WVPSPASRPEVPTAEKDL-----SSKGKVS-----DLSLGSVCCKHKPSSSTLSL 316  
 247 YLKERRKQKRVNPKVDCNAPCITPSTNGKVSTGNDENMLHIGINGSL-----TDFV 298  
 317 SPPRQNAFOLRPFIETRHFLYSRGDGOERLNPSFLNSLQPNLTGARRLVEIF----- 370  
 299 K-----QAKQVRKNKVPFGLSETYSVIPPNHILKTLRPNCSDSLMMHIFGEVNVW 351  
 371 -----LGSRPRTSGPLCRTHLSRRYQMRPLFOQLLVNHAECQVRELLRSHC----- 418  
 352 STTFSHGKGNCP--SGSICDIYHSLK---SLKMLIGTKSSHLK-----MLDKHCPVLL 402  
 419 -----RFTANQOVTDALNTSPPH-----LMDLLRLHSSPWQVYGFRL 456  
 403 QEDALKSGTTSQSSRRQKADKL---PHGSSSSQTKPKPCPSVEERKLYCTNDQVVSPFW 458  
 457 ACLCKVYSASLWGTNRNERFFKNLKKFISLKGKLSLOELMWKM 506  
 459 AICRYIIPESLLGTTHQMRVLRKXNIAWFSRRRNEKCTVNOFLHKVYKPSDFPFPAKELC 518  
 507 C---HWLRSSPGKDRVPAAEHRLRERILAT-FLFWLMDTYVYVQLLRSPFYITESTFOKN 561  
 519 CMVNGHELOSE-----SIRSTQCMCTKWSLFWLEIVKVLVHFNFYATESQGRL 569  
 562 RLPFYKSVWSKJQSIGVROHLRERVLRELSEBEVHHQDTWLAMPICRLRIPKPNGLR 621  
 570 NIYYRKRSWERLISKEISKALDGYVLVDAAEASRRK-----LSKRFRLPKANGVR 622  
 622 PIVNYSYMGTRALGRKQAOHTORLKTILFSLMNYERTKHPHMGSSVLGMNDIYRTWR 681  
 623 MVLDFSSSR-----SQSLRDTHAVLKDIQKEPDVLGSSVDFDHDYRNLK 669  
 682 AFLVLRBALD-QTPRMVYFKADVTGAYDAIPQGLKEVVVANMRHSESTYCIQYAVV-- 738  
 670 PYLIHLRSQSGELPPLYFYVADVFKAFDSVDQGLLHVIOFLK---DEVILNRCRLVCC 726  
 739 -RRDS-----QGVHKSFRQVTTLSLQVPMGQFLKHLQDSDASALRNSVITEQIS 790  
 727 GKRSNNWNKILVSSDKNSNFSRFTSV---PY-----NALQ-SIVVDKGN 768  
 791 MNESSSLFDFFLHFLRHSVVTGDCRCYTCCQIGIPQGSLSLTLLCSLCFGDMENKL---- 846  
 769 HRVKKDLMWIGNMLKNNMLQDKSFYQIAGIPQGHRLSSILCCFYFGHLERTLLIYFF 828  
 847 PAEVQORD-----GLLLRFVDDFLVTPHLDQAKTFLSTLVHGVPEYGC 889  
 829 LEEASKDVSSKECSREELIPTSXYKLRFIDYDLFVSTSRDQASFYHRLKHGFKDYN 888  
 890 MINLQKTVNFFPVPGLGGAAPYQPAHCLF-----PWCGLLDTOTLEVFCDYS 940  
 889 FMNETKCFINFE-----DKBEHRCSSNRMFVGNGVFPVWVTGLLINSRTFEVQVY 941  
 941 GYATSIKTSLTFTQSVFKAGKTWRNKLKLSVLRLKCHGLFLDQVNSLQTVCIYKIFLL 1000  
 942 RYLSGHISSTFSVAVQWQKPVNLRQKLCYFLVPKHPILFDSNINSGEIVRLNIYQIFLL 1001  
 1001 QAYEFHACVQLPFDQVRKNTLF-FLGIISQASCCYAILKVKPNPGMTLKASGSPFP-- 1057  
 1002 AAMKFCYVTEVERFWKHLHPOTLKFITI-----SVYMFRLINRRVRINTSGSFRPVL 1056  
 1058 ----EAAHW-CYQAFLLKLAHSVYIKCLLGLPLRTAQKLCRLKLPENTWILKAADPAL 1113  
 1057 KLYKEEVIWGLDAYIQVLKKNRVRMELLIVLXSA--LSKHSLSQLSSELKLYATDRSN 1114  
 1114 ST 1115  
 1115 SS 1116

```
352 STTPSHKGNCP--SGSICLYHSLK---SLKNLIGTKSSHLK-----MLLDKHPVLL 402
419 -----RFTANQVTDALNTSPH-----LMDLLRLHSSPWQVYGLR 456
403 QEDALKSGTTSQSSRRQKADKL---PHGSSSSQTKGPKCPSVEERKLYCTNDQVFW 458
457 ACLCKVVSASLWGTNRHNRPPKLNKKFISLKGKYSLOELMMKMKVED----- 506
459 AICRYIVPESILGTTTHQVRLKNTAFVSRREKCTVNGFLHKVPSDFPFARKELC 518
507 C-----HMLRSSPKORVPAEHLREILAT-FLWMDTVVQLLSFFVITESTQKN 561
519 CMVNGHELQSE-----SIRSTQMLCTKWISLWFLKLVHFNFYATESQGR 569
562 RLFFYKSVKSLQSIGVROHLERVLRELSEVRRHODTWLAMPICRLRFIPKPNGLR 621
570 NIYYRKRWERLISKEISKALDGVYVDDAEASSRK-----LSKFLPKANGVR 622
622 PIVNMSYSGTRALGRKQAOHFTQRLKTLFSLMLNYETKPHLMGSSVLGMNDIYRWR 681
623 MWLDFSSSR-----SQSLRDTHAVLKDIQKEPDLVGLSSVPDHDFFYNLC 669
682 AFVLRVRLD-QTPMYFVKADVTGAYDAIPOGKLVEVYVANKIRHSESTYCIQYAVV-- 738
670 PYLHLRSQSGELPPLYFYVADVFKAFDSVQGLLHVIOQSLK---DEYILNRLVCC 726
739 -RRDS-----QGQVHKSFRQVTTLSDLQPMQOFLKHLQDSASALRNSVIEOSIS 790
727 GKSNWVKILYSSDKNSFSFTSTV---PY-----NALQ-SIVVDKGEN 768
791 MNSSSLDFDFLHFLRHSVVKIGDRCTYQCOGIPQSSSLTLLCSLCFGDMENKL----- 846
769 HRYRKDLMVWIGNMLKNNMLQDKSFYVQIAGIPQGHRLSLSLCCFYGHLETLIYFF 828
847 FAZVQD-----GULLRFVDDFLVTPHLDQAKTFLSLTVHGVPEYGC 889
829 LEASKDVSKESCEEBELIIPTSYKLLRFIDYLVFVSTRDQASSFYHRLKSGFKDYN 888
890 MINLQKTVNVFVPEOTLGGAAPOQLPAHCLF-----PWCGLLDTOTLEVCDS 940
889 FMNETKFCINFE-----DKEHRCSNRMFVGONGVFPVVRWTGLLINSRTFEVDYT 941
941 GYAQTSIKTSLTFQSVFKAGTMRNKLVLKCHGLFLDLQVNSLQTVGINIYKIFLL 1000
942 RYLSGHISSTFSVAMQKPVRLRQKCYFLVPKCHPILFDSNINSGEIVRLNIYQIFLL 1001
1001 QAYRFACVQLPFDQVRKNLTF-FLGISSQASCYAILKVNPGMTLKASGSPFP-- 1057
1002 AAMKFCYVYVSRFWKLPQTLFKFITI-----SVRYMFLINRRVRRINTGSSFRVL 1056
1058 -----EAAHWLCYQAFLLKLAHSAVYKLLGRLTAQKLLCEKLPBETWTLKAAADPAL 1113
1057 KLYKEEVIWGLDAYIQVLKKNRYRMLLIYKGA--LSKHSLSQQLSSELRYATDRSN 1114
1114 ST 1115
1115 SS 1116

RESULT 12
Q9AU13 PRELIMINARY; PRT; 1259 AA.
AC Q9AU13;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Telomerase reverse transcriptase.
GN TEST.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
```

```

[1]
SEQUENCE FROM N.A.
Oguchi K., Tamura K., Takahashi H.;
"Molecular cloning and characterization of OSTERT, a telomerase
reverse transcriptase homolog in Oryza sativa.";
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
EMBL: AF288216; AAK35007.1; -
Gramene; Q9AU13; -
InterPro; IPR001209; Ribosomal_S14.
InterPro; IPR004077; RVtse.
InterPro; IPR003545; Telomerase_RT.
Pfam; PF000078; rvt; 2.
PRINTS; PR01365; TELOMERASERT.
PROSITE; PS00527; RIBOSOMAL_S14; 1.
RNA-directed DNA polymerase; transferase.
SEQUENCE 1259 AA; 143710 MW; 15B041789F2D5CAD CRC64;

Query Match      12.4%; Score 729; DB 10; Length 1259;
Best local Similarity 24.2%; Pred. No. 2.7e-52;
Matches 297; Conservative 182; Mismatches 429; Indels 318; Gaps 47;

      4 APR-----CPAVERLLRGRYREVWPLATFVRRLGPEGRRILVQGDPIKYRTLV--AQCIV 56
      |||
      41 SPRHCSPCAPCRGVAG-----CLACPR-----WEHLRGGDIATRRILITRAVCAT 89
      |||
      57 CMHWGQPP-----PADLSFHQVSLKELVARVQRLCERNERNVLAFGFELLNEARGGP 111
      |||
      90 AADDIASAPPPRYTPGNSGHGQARLVREMKSIIVADQ-SHGTKNVLCNLH-----EGQG 143
      |||
      112 PMAFTSSVRSYLPNTVIETLVASGAWMLLSRVGDDLVLVLAHFCALYLLVPPSCAYQVC 171
      |||
      144 STC-----ISDLVSSSWILLHRIGDLLMVCILRCTSIFLPKKNIDYFOVS 190
      |||
      172 GSPLYQI-----CATTDIWPS-----VSASY 192
      |||
      191 GYPLNVLIRNPIFASTVARKHPQPTTKAKCHTCYLVKASNAENLSICHDSNSGNWSSF 250
      |||
      193 RTRPV--GRNFTNLRFFQQIKSSRQEAQPLALPSG-----TKRHLSITS 238
      |||
      251 STCKIVTQSCETCGSIRRAESKDPSEGCNCPKPPDSGSECCNCTHTNTRKKRLYS 310
      |||
      239 TGVPSAKKARCYVPRVE-----GPHRQVLTPSGKSWYFSPARSPE 281
      |||
      311 WQRRSKKIQVC---SVDESSAEWSKLGNSFNFMNGPSNL---AGK--MNDQAQVSE 360
      |||
      282 VPT-----AEKDLSSKGYSD--LSLGSVC-----CKHKPSTSLTSP 319
      |||
      361 LTVDNTSLARNDDSSSIKVINATILSSKSPCVDFIRGSGLSCHYLSSEVOYQSTC 420
      |||
      320 RQ-----NA--FOLRPFIETRHFLYSRGDGQRLNFPSPLLS-N 354
      |||
      421 PQWGPSSYLHLNSCSICFNCIIISNAKSLSDLSLRNGIFYNRRITVSVFCHKILSKR 480
      |||
      355 LQPNLTGARLVEIIF-----LGRPRTSQPLCETHRLSRYHQMPLEFQQ 400
      |||
      481 KRPD---ALSLVKHIFGINSCCASLLKYNCHESITIRKSNCLC-----CWLPKSI-KN 528
      |||
      401 LLVNHAECQYVRLLRSHCRPR-----TANQQVTALNTSP-----HLMDLLRLH 445
      |||
      529 LIIRNKRQYKKLFLKHSVVKCVADPTKN---DGKAHYPPGKAAYDRSFSRLBAY 584
      |||
      445 SSPWQVYGFACLCKVVSASLWQTRNRRRFPNKKFISLQKYGKLSQELMMWKVE 505
      |||
      585 STHQVQAVFWAVLKRIVPKPELLNGFFGKESLRATNIWKFLKRRFETPOLSDICGLKVS 644
      |||
      506 DCHWLRSSP-----GKD--RVPAASHRLRERILATFLFWMMDTVVOLLRSFPVI 553
      |||
      645 HYSWLSNIEFSNCFCSAIIHQTGSSTSABEQKQLHCWISWLFSDVIVPVRTYFV 704
      |||
      554 TESTTQKNRLFFPKYSVWSKLOSTGVQHLRLVRLRELQOE---EVRHHQDTWLAMPICR 610
      |||
      705 TERESKRYDVFYPKSVWRDITSNAIAS-LFNKNFRLRGPRAKVRHLNCS-----SR 757

```

Qy	611	LRF	PKNGNLRPIVNMSSYSGTRALGRKQAOHQFTORLKT	LF	SMIN	YRTKPHLMGSSV	670
Db	758	VRFLPKAKMRPLVDL	-----RAGSKDANLNKCHLIMKKLRDEKPEMPGSSV	804			
Qy	671	LGNMDIVRTWRAFLVRVA	-LDQTPRMFVKADVTGAYDAI	POGKLVEVVANMIRHSEST	729		
Db	805	FDYNNVHQNLQSFISSKRSQMKKLKVIIVADVSKAFDCVSHDWMLKMLDDDAFKDE	---	862			
Qy	730	YCIHQYA-VVRRDSQGOVHKHSFRFRQVTTLSDLQPYMGQFLKHLQDSASALRNSVYIEQS	788				
Db	863	YTVRKCSKVLCNRKSNLYRF	-----DSNAGISGNSIYDLS	899			
Qy	789	ISXNNESSSLF	-----DFFLHFL	-----RHSVWKIGDECIYQCQGIPOGSSLSL	833		
Db	900	IQUS-SGGGIFVQGTICRLKEQFHLLYEQIKCNILKIGQKYIQQVGIAQGSKLSPN	958				
Qy	834	LCSLCFGDMENKLF	AEVORDG-----LLRFVDDFLLVTPHLDQAKTFLSTL	880			
Db	959	LCSLYGHLENSVLKFLHDSKLNAGEATSEPEYLLMR	FIDDFIFISFSLHAQKFLNRM	1018			
Qy	881	VHGVEPGCMINLOKTVVNPVPEFTGLGNA	PYOLPAH-----CLFPWCGLLLDDTQTL	934			
Db	1019	RRGFVFVNCYMDSKYGFN	-----CAGNSEPSNLYRGDGVSPMPWSGLINCETLE	1073			
Qy	935	VPCDYSYACTS	IKTSLTFSQVFKAGTWRNKLLSVLRKCHGLFEDLVNSLQTVGINI	994			
Db	1074	IQADYRFLDITIIISITVK-MHSSYKIYHSLKCHYMRKACHPI	FIFDSDNSINSPGTRVNI	1132			
Qy	995	YKIFLQAVRFHACVTL	PFDPQVRK	1020			
Db	1133	YQAFLLCAMKFH-CYIRSVSDANVSK	1157				

RESULT 13

Q8LKWQ PRELIMINARY; PRT; 1261 AA.

ID

Q8LKWQ

PRELIMINARY;

PRT;

1261

AA.

AC

Q8LKWQ;

DT

01-OCT-2002 (TrEMBLrel. 22, Created)

DT

01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT

01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE

Telomerase reverse transcriptase.

GN

TERF.

OS

Oryza sativa ('japonica cultivar-group').

OC

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC

Ehrhartoideae; Oryzeae; Oryza.

OX

NCBI\_TaxID=39947;

RN

[1]

RP

SEQUENCE FROM N.A.

RC

STRAIN=cv. Nipponbare;

RA

MEDLINE=22096152; PubMed=12100484;

RX

Heller-Uzysynska K., Schnippenkoetter W., Kilian A.;

RT

"Cloning and characterization of rice (Oryza sativa L) telomerase

RT

reverse transcriptase, which reveals complex splicing patterns.";

EL

Plant J 31:75-86(2002).

DR

EMBL; AF494453; AAM21641.1; -.

DR

Gramene; Q8LKWQ; -.

DR

InterPro; IPR001209; Ribosomal\_S14.

DR

InterPro; IPR000477; RVtsel.

DR

InterPro; IPR003545; Telomerase\_RT.

DR

Pfam; PF000078; rvt; 2

DR

PRINTS; PR01365; TELOMERASERT.

DR

PROSITE; PS00527; RIBOSOMAL\_S14; 1.

KW

RNA-directed DNA polymerase; Transferase.

SQ

SEQUENCE 1261 AA; 143715 MW; 353153D1C8B7D5CB CRC64;

Query Match

11.6%; Score 686; DB 10; Length 1261;

Best local similarity

23.9%; Pred. No. 1.2e-48;

Matches 294; Conservative 181; Mismatches 432; Indels 322; Gaps 49;

Qy	4	APR	-----CPAVRSLRLRSYREVNPLATFVRLGPEGRRLVQGFQKIYRTLV	-----AQCIV	56
Db	41	SPRHCBSPCACRGRVAGS	-----CLACRR	-----WEHLRDGDFVAYRRLITRAVCAI	89



57 CMWGSOPP-----PADLSFHOVSSSLKELVARVQVORLCERNERNVLAFGFELLNEARGP 111  
90 AADDLSAPPPPPPTPGNSHGSCAELVEMKXSIADQ-SHGTKNVLCNGUH-----EGGQ 143  
112 PMAFTSSVRSYLPNTVITETLRVSGAMWLLSRVGGDLDLVYLLAHACALYLLPPSCAYQVC 171  
144 SIC-----ISDLVSSSSWSIILHRIGDLMLCYLLRCSITIFLPVKNDYFQVS 190  
172 GSPLYQI-----CATTDIWP-----VSASY 192  
191 GVPLNVLRNFIFASTVARKHQPTTKAKCTCYLWKSANMAENLSICHDSNSGVNSSF 250  
193 RPTPRV-GRNFTNLRFLQOKSSROEAPKPLALPSRG-----TKRHLSTLS 238  
251 SSTKIVTQOCETCGSIRRAESKDPGECNCPKPSDGRSGCCNCCYTHNTRKRLYS 310  
239 TSVPASAKACYPVPRVE-----GPHRQVLPFSGSKSWPSPARSP 281  
311 WORRSKKQVC-----SVDESSAEMSKLNGSNFMNSGPNL-----AGK--MNDQAQSYE 360  
282 VPT-----AEKDLSSKGVSD--LSLSGVC-----CKKPSSTLSLSP 319  
361 LTVDNVTLARNVDSSSEIKVINATILSSEKSPCSVFDIRGSOGLSCHYSLSVEVQVQSTC 420  
320 RQ-----NA-----FOLRPIETRFYLSRGDQGRNLPSPFLLS-N 354  
421 PQVGESSYLHNSCISCFNCIISNASKHLSLDSLSIRNGIFYNRRRTYSVFHCKHLSKR 480  
355 LOPNLGTARRVEIIF-----LGRPRTSGPLCRLTHLSRRYQMRPLFOQ 400  
481 KPDP-----ALSIVKHIFGINSOCASLLKYNCHSTIRKSNCLC-----CWLPKSI-KN 528  
401 LLVNHAECQYVRLRSHCRFR-----TANQQVTDALNTSP-----HLMDLRLRH 445  
529 LIRNSKQYKXKFLKHGCVKCKVADPTKN-----DGKAHYPPGKAAYDRSFSRLAY 584  
446 SSPWQVYGLFACLVKYSASLWGTRENERFFKQLKFKISLKGKLSIQELMWKQVYE 505  
585 SHQOVASVFWAVLKRIVPKPLNGSGFKSLRNTIWKFLRRFETFLQSDCIGDLKVS 644  
506 DCHWLRSSP-----GKD--RVFAAHRERILATFLFWMMDTVYVQLRFFYI 553  
645 HYSWLSNIEFNSCFCSAIIQGTGSSISAEQKQNLHOMISWLPDSDIVPVRTYFV 704  
554 TESTQKRLFPYKSVWSKLSQISQVQHLRVLRELQSE-----EVRHODTWLAMPICR 610  
705 TERESKRYDVFPYKSVWRDLTSNAIAS-LNKKCNFRILRGEPRKAVRLNCS-----SR 757  
611 LRFIPKPNGLRPIVNMYSYMGTRALGRRKQAOHFTQRLKTLFSLMNLVYRTKHPHLMGSSV 670  
758 VPFLPKADMRPLVDL-----RAKSKDANLKNCHILMKLXDEKPEMGSSV 804  
671 LGMNDIYTRAFVLRVRA-IDQTPRYFVKADVTGAYDAIPQGLVEVVANMIRHSEST 729  
805 FDNVNHQNLQFISSEKASQLMKKLKVIYVADVSKAFDCVSHDMVLMKIDDAFKODE-- 862  
730 YCIRQYA-VVRDSOGVHKSFRRQVITSLDLPYMGQFLKHLDQSDASALRNSVVEIOS 788  
863 YTVKCKSVKVNRSKNSLYR-----DSNAGSNGNSLYDS 899  
789 ISMNESSSLF-----DFFLHFL-----RHSVVKIGDRCYTCQCGIPQSSLSLT 833  
900 IQLS-SGGGIFVDQGTICRLKEQFHHLLVBOIKCNILKIGQYKYLQVQVGAQSKLSPN 958  
834 LCSLCFGDMENKLAERVORDG-----LLLRFDVDDLLVTPHLDQAKTFLSTL 880  
959 LCSLYYGHLENSVLSKFLHDSKLNAGEAFSPFYLMLRFIDDFIFISFSLHQAQKFLNRM 1018  
881 VHGVPEYGCMLNLOKTVVNFVPEPQTGGAPYQLPAH-----CLPPWCGLLLDQTLE 934  
1019 REGFVFNMYNDSKYGNF-----CAGNSEPSSNRLYRGDGDGVFWPNSGLLINCETLE 1073

935 VFDYSGYAQTS--IKTSLTFSQVFKAGTKMKNLLSVLRK-CHGLFELDLQVNSLQTV 991  
1074 IQADYTRYDITHYLPVSLVRLCHLQSI--SKLCLYAVQNAHPFYDSDNINSPTGIR 1131  
992 INYKIFLQVAFRHACVITQLPFDQVRK 1020  
1132 VNITYQAFLLCANKFH-CYIRSVSDANVSK 1159

RESULT 14  
Q9JLMI PRELIMINARY; PRT; 104 AA.  
AC Q9JLMI;  
DT 01-OCT-2000 (T-EMBLrel. 15, Created)  
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)  
DE Telomerase reverse transcriptase (Fragment).  
GN TERT.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Hatakeyama S., Ishikawa F.;  
RT "Identification of the mouse telomerase reverse transcriptase (mTERT)  
promoter.";  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF157502; AAP42984.1; --  
DR MGD; MGI:1202709; TERT.  
KW RNA-directed DNA polymerase.  
FT NON TER 104 104  
SQ SEQUENCE 104 AA; 11998 MW; 4B649B63476D3D44 CRC64;

Query Match 9.3%; Score 550; DB 11; Length 104;  
Best Local Similarity 100.0%; Pred. No. 9.9e-39;  
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRAPCPAIVRSYREVWPLATFVRLPGEGRRLVQGPDKPIYETLVAQCILVCMHW 60  
DB 1 MTRAPCPAIVRSYREVWPLATFVRLPGEGRRLVQGPDKPIYETLVAQCILVCMHW 60  
QY 61 GSOPPPADLSFHQVSSSLKELVARVQVORLCERNERNVLAFGFELL 104  
DB 61 GSOPPPADLSFHQVSSSLKELVARVQVORLCERNERNVLAFGFELL 104

RESULT 15  
Q8MUB3 PRELIMINARY; PRT; 1032 AA.  
AC Q8MUB3;  
DT 01-OCT-2002 (T-EMBLrel. 22, Created)  
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
DE Telomerase reverse transcriptase.  
GN TERT-1.  
OS Euplotes crassus.  
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;  
OC Euploidea; Euplotidae; Moneuplotes.  
OX NCBI\_TaxID=5936;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang L., Dean S.R., Shippen D.E.;  
RT "Oligomerization of the Telomerase Reverse Transcriptase from Euplotes  
crassus.";  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF528527; AM95622.1; --  
DR InterPro; IPR000477; RVTse.  
DR InterPro; IPR003545; Telomerase\_RT.  
DR Pfam; PF00078; rvt; 2.  
DR PRINTS; PR01365; TELOMERASERT.  
KW RNA-directed DNA polymerase; Transferase.  
SQ SEQUENCE 1032 AA; 122966 MW; 0C401515839801A8 CRC64;

```

Query Match          7.6%; Score 450; DB 5; Length 1032;
Best Local Similarity 22.7%; Pred No. 8.6e-29;
Matches 179; Conservative 142; Mismatches 316; Indels 150; Gaps 32;

392 WQWRPLFQQLLVNH-----AECQYVLLRSHCRPTANQOQVTDALN-----TSPPHLMD 440
323 WVER- --KVILSHLNDRAKAKAYEELFS--YTIENKCVTOFLNEFFYHTLPKDFM- 375
441 LLRLHSSPWQVYGLFACLVCKVVSASLQWTRNRRERFFKNLKPISLKGKGLSLQELMW 500
376 -----TGKRSFQSKIKYVGLNKHLEIHKNLILG 406
501 KWKVEDCHWLRRSPGK-----DRVPAEHRRLERILATFLWMDTYVQVLLRSFYIT 554
407 KLDIKSIKWTGFKTSKNVHYFDK-----ENRF--VLWRILRWIPEDVVVSLIRCFYVT 459
555 ESTFQKNRLLFFYKSVMSKLSIGVROHLERVRLRELSEEVRRHODTWLAMPICRLRFI 614
460 EQOKSYSETYYIKNTWIVMKYSIAD-LNRETLEEVPEKEDWVKWKNELRFAP-GKLRLI 517
615 PKPGLRPIVYMSYMGTRALGRKQAOHFTQRLKTLPSMLNYERTKPHL---MGSSVL 671
518 PKKTTFRPIN---TFNKILDPDGKTSQMTTNTKLSYSHLMLKTLKRMFKDPFGPAVF 573
672 GMDIVHTWEAFVLRVRLDQTPRMVYFKADVTGAYDAIPQCKLVEVVANMIRHSESTYC 731
574 NYDDVMKYEFEFVLRWQVGR-FGLYFTVMDIEKYDSVDRKLSQLLGT-TRLLSEFR 631
732 IRQYAVVRD-----SQQVHKSPR---ROVTTLSLQPYMGQFLXHQD--SDA 776
632 IMTVQAMKRNNAEVVDLDKCAKNKKECFKQWKIALEGDQYP---SLNVLEDQNDL 688
777 SALENSVVIESLSMNESSSLFDFLHRLSHSVKIGDRCYTOCGIPQGSSTLLCS 836
689 NA-KETLLVENQRDPYKXALLDPVIKICRNYIEFRKYKQTKGIPQGLCVSSILSS 747
837 LCFGDMENKLFVAVQRDQ-----LLLRVDDPFLVTPHLDQAKTFLSTLVHGVPEY 887
748 FYASLEENALGYLRKESMDANDPNTLLMLRTDDYLLITTKENNAIFIEKLINVRQN 807
888 GCMINLQKTVNRPVPGTLG---GAAPYOLPAHCLPWCGLLDTOTLE----- 934
808 KFKENMKKLOTNPFDPFSKLNKYGMASVEDQNIADHYIDWIGISIDMTALMPNINLR 867
935 --VFCDYSGYAQTSIKTSLTFQSVFVAGTKMKNKLSVLRKCHGLFLDLQV---SLQT 989
868 KGILCTLNMNQTK-----KASWMLKRLKLSFLMNNITHYPRKTTITNREFSNKT 916
990 VCINIYKIFLLQAYRPHACVQLPPDQVRKNL-----TFFGLISS--QASCCYAIL 1040
917 ----LNKLYIAGAYKNQCCIE--YKDHFKTNTIHPQLDKIICAIYISVTRAFKYLVC 970
1041 KVKNPGMTLKASGSFP-PEAAHWLCYQAFLLKLAHRSVIYKCLGLPLRTAQKILCKLPE 1099
971 NIKDSIYDREHSRDFSIALRPHI--EIFSLKKNHFGVVKILVA---KEQKMLEEK-DN 1024
1100 ATWTILK 1106
1025 ACMTAIE 1031

```

rch completed: February 4, 2004, 14:11:48  
time : 50 secs